PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| (51) International Patent Classification 6: | | (11) International Publication Number: | WO 99/29870 |
|---|----|--|-------------------------|
| C12N 15/31, C07K 14/195, A61K 38/00, 38/16, 39/00 | A1 | (43) International Publication Date: | 17 June 1999 (17.06.99) |
| | | | |

PCT/AU98/01023 (21) International Application Number:

10 December 1998 (10.12.98) (22) International Filing Date:

(30) Priority Data:

| 10 December 1997 (10.12.97) | ΑU |
|-----------------------------|---|
| 31 December 1997 (31.12.97) | ΑU |
| 30 January 1998 (30.01.98) | ΑU |
| 10 March 1998 (10.03.98) | ΑU |
| 9 April 1998 (09.04.98) | ΑU |
| 23 April 1998 (23.04.98) | ΑU |
| 5 May 1998 (05.05.98) | ΑU |
| 22 May 1998 (22.05.98) | ΑU |
| 29 July 1998 (29.07.98) | ΑU |
| 30 July 1998 (30.07.98) | ΑU |
| 4 August 1998 (04.08.98) | ΑU |
| | 31 December 1997 (31.12.97) 30 January 1998 (30.01.98) 10 March 1998 (10.03.98) 9 April 1998 (09.04.98) 23 April 1998 (23.04.98) 5 May 1998 (05.05.98) 22 May 1998 (22.05.98) 29 July 1998 (29.07.98) 30 July 1998 (30.07.98) |

(71) Applicant (for all designated States except US): CSL LIMITED [AU/AU]; 45 Poplar Road, Parkville, VIC 3052 (AU).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ROSS, Bruce, Carter [AU/AU]; 32 Kelson Street, Coburg, VIC 3058 (AU). BARR, Ian, George [AU/AU]; 13 Shakespeare Drive, Templestowe, VIC 3106 (AU). PATTERSON, Michelle, Anne [AU/AU]; 14 Cotterell Way, Laverton, VIC 3028

(AU). AGIUS, Catherine, Therese [AU/AU]; 250 Elgar Road, Box Hill South, VIC 3128 (AU). ROTHEL, Linda, Joy [AU/AU]; 10 Rothschild Street, Glen Huntly, VIC 3163 (AU). MARGETTS, Mai, Brigid [IE/AU]; 92 Bent Street, Moonee Ponds, VIC 3039 (AU). HOCKING, Dianna, Margaret [AU/AU]; 49 Illawarra Road, Flemington, VIC 3031 (AU). WEBB, Elizabeth, Ann [AU/AU]; 36 Zigzag Road, Eltham, VIC 3422 (AU).

(74) Agent: F.B. RICE & CO.; 605 Darling Street, Balmain, NSW 2041 (AU).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: PORPHORYMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES

(57) Abstract

The present invention relates to isolated Porphorymonas gingivalis polypeptides and nucleotides. The polypeptides include: an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or an amino acid sequence at least 85 %, preferably at least 95 %, identical to an amino acid sequence selected from the group consisting of SEO. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
|----|--------------------------|-----|---------------------|----------|-----------------------|----|-----------------------------|
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | Republic of Macedonia | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | • |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Trinidad and Tobago Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | |
| BY | Belarus | IS | Iceland | MW | Malawi | | Uganda |
| CA | Canada | IT | Italy | MX | Mexico | US | United States of America |
| CF | Central African Republic | JР | Japan | NE. | Niger | UZ | Uzbekistan |
| CG | Congo | KE | Kenya | NL NL | Netherlands | VN | Viet Nam |
| CH | Switzerland | KG | Kyrgyzstan | NO NO | | YU | Yugoslavia |
| CI | Côte d'Ivoire | KP | Democratic People's | | Norway | zw | Zimbabwe |
| CM | Cameroon | 141 | Republic of Korea | NZ | New Zealand | | |
| CN | China | KR | Republic of Korea | PL | Poland | | |
| CU | Cuba | KZ | Kazakstan | PT | Portugal | | |
| CZ | Czech Republic | LC | Saint Lucia | RO | Romania | | |
| DE | Germany | Li | Liechtenstein | RU | Russian Federation | | |
| DK | Denmark | LK | | SD | Sudan | | |
| EE | Estonia | | Sri Lanka | SE | Sweden | | |
| | Latoma | LR | Liberia | SG | Singapore | | |
| | | | | | | | |

Porphorymonas gingivalis polypeptides and nucleotides

FIELD OF THE INVENTION

5

10

15

20

25

30

35

The present invention relates to *P. gingivalis* nucleotide sequences, *P. gingivalis* polypeptides and probes for detection of *P. gingivalis*. The *P. gingivalis* polypeptides and nucleotides can be used in compositions for use in raising an immune response in a subject against *P. gingivalis* and treating or preventing or reducing the severity of the condition known as periodontitis.

BACKGROUND OF THE INVENTION

Periodontal diseases are bacterial-associated inflammatory diseases of the supporting tissues of the teeth and range from the relatively mild form of gingivitis, the non-specific, reversible inflammation of gingival tissue to the more aggressive forms of periodontitis which are characterised by the destruction of the tooth's supporting structures. Periodontitis is associated with a subgingival infection of a consortium of specific Gram-negative bacteria that leads to the destruction of the periodontium and is a major public health problem. One bacterium that has attracted considerable interest is P. gingivalis as the recovery of this microorganism from adult periodontitis lesions can be up to 50% of the subgingival anaerobically cultivable flora, whereas P. gingivalis is rarely recovered, and then in low numbers, from healthy sites. A proportional increase in the level of P. gingivalis in subgingival plaque has been associated with an increased severity of periodontitis and eradication of the microorganism from the cultivable subgingival microbial population is accompanied by resolution of the disease. The progression of periodontitis lesions in non-human primates has been demonstrated with the subgingival implantation of P. gingivalis. These findings in both animals and humans suggest a major role for P. gingivalis in the development of adult periodontitis.

P. gingivalis is a black-pigmented, anaerobic, asaccharolytic, proteolytic Gram-negative rod that obtains energy from the metabolism of specific amino acids. The microorganism has an absolute growth requirement for iron, preferentially in the form of haeme or its Fe(III)

10

15

20

25

30

oxidation product haemin and when grown under conditions of excess haemin is highly virulent in experimental animals. A number of virulence factors have been implicated in the pathogenicity of *P. gingivalis* including the capsule, adhesins, cytotoxins and extracellular hydrolytic enzymes.

In order to develop an efficacious and safe vaccine to prevent, eliminate or reduce P. gingivalis colonisation it is necessary to identify and produce antigens that are involved in virulence that have utility as immunogens possibly through the generation of specific antibodies. Whilst it is possible to attempt to isolate antigens directly from cultures of P. gingivalis this is often difficult. For example as mentioned above, P. gingivalis is a strict anaerobe and can be difficult to isolate and grow. It is also known that, for a number of organisms, when cultured in vitro that many virulence genes are down regulated and the encoded proteins are no longer expressed. If conventional chemistry techniques were applied to purify vaccine candidates potentially important (protective) molecules may not be identified. With DNA sequencing, as the gene is present (but not transcribed) even when the organism is grown in vitro it can be identified, cloned and produced as a recombinant DNA protein. Similarly, a protective antigen or therapeutic target may be transiently expressed by the organism in vitro or produced in low levels making the identification of these molecules extremely difficult by conventional methods.

With serological identification of therapeutic targets one is limited to those responses which are detectable using standard methods such as Western Blotting or ELISA. The limitation here is the both the level of response that is generated by the animal or human and determining whether this response is protective, damaging or irrelevant. No such limitation is present with a sequencing approach to the identification of potential therapeutic or prophylactic targets.

It is also well known that *P. gingivalis* produces a range of broadly active proteases (University of Melbourne International Patent Application No PCT /AU 96/00673, US Patent Nos 5,475,097 and 5,523,390), which make the identification of intact proteins difficult because of their degradation by these proteases.

10

15

20

25

30

35

SUMMARY OF THE INVENTION

The present inventors have attempted to isolate P. gingivalis nucleotide sequences which can be used for recombinant production of P. gingivalis polypeptides and to develop nucleotide probes specific for P. gingivalis. The DNA sequences listed below have been selected from a large number of P. gingivalis sequences according to their indicative potential as vaccine candidates. This intuitive step involved comparison of the deduced protein sequence from the P. gingivalis DNA sequences to the known protein sequence databases. Some of the characteristics used to select useful vaccine candidates include; the expected cellular location, such as outer membrane proteins or secreted proteins, particular functional activities of similar proteins such as those with an enzymatic or proteolytic activity, proteins involved in essential metabolic pathways that when inactivated or blocked may be deleterious or lethal to the organism, proteins that might be expected to play a role in the pathogenesis of the organism eg. red cell lysis, cell agglutination or cell receptors and proteins which are paralogues to proteins with proven vaccine efficacy.

In a first aspect the present invention consists an isolated antigenic *Porphorymonas gingivalis* polypeptide, the polypeptide comprising;

an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or

an amino acid sequence at least 85%, preferably at least 95%, identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or

at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532.

In an embodiment of the present invention the polypeptide comprises;

an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532; or

4

an amino acid sequence at least 85%, preferably at least 95%, identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532; or

at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532.

As used herein % identity for polypeptides is to be calculated using the alignment algorithm of Needleman and Munsch (9) using a standard protein scoring matrix (Blosum 50).

In a preferred embodiment of the present invention the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386, SEQ. ID. NO. 424, SEQ. ID. NO. 425, SEQ. ID. NO. 434, SEQ. ID. NO. 447, SEQ. ID. NO. 458, SEQ. ID. NO. 475, SEQ. ID. NO. 498, 15 SEQ. ID. NO. 499, SEQ. ID. NO. 500, SEQ. ID. NO. 501, SEQ. ID. NO. 387, SEQ. ID. NO. 400, SEQ. ID. NO. 411, SEQ. ID. NO. 419, SEQ. ID. NO. 420, SEQ. ID. NO. 427, SEQ. ID. NO. 429, SEQ. ID. NO. 433, SEQ. ID. NO. 437, SEQ. ID. NO. 438, SEQ. ID. NO. 443, SEQ. ID. NO. 444, SEQ. ID. NO. 448, 20 SEQ. ID. NO. 449, SEQ. ID. NO. 452, SEQ. ID. NO. 455, SEQ. ID. NO. 457, SEQ. ID. NO. 459, SEQ. ID. NO. 461, SEQ. ID. NO. 462, SEQ. ID. NO. 463, SEQ. ID. NO. 467, SEQ. ID. NO. 468, SEQ. ID. NO. 469, SEQ. ID. NO. 482, SEQ. ID. NO. 484, SEQ. ID. NO. 485, SEQ. ID. NO. 494, SEQ. ID. NO. 508, SEQ. ID. NO. 509, SEQ. ID. NO. 510, SEQ. ID. NO. 520, SEQ. ID. NO. 521, 25 SEQ. ID. NO. 522, SEQ. ID. NO. 525, SEQ. ID. NO. 526, SEQ. ID. NO. 528, SEQ. ID. NO. 389, SEQ. ID. NO. 390 and SEQ. ID. NO. 391.

In another preferred embodiment of the present invention the polypeptide comprises an amino acid sequence selected from the group consisting of residue 422 to residue 531 of SEQ. ID. NO. 303, residue 534 to residue 582 of SEQ. ID. NO. 303, residue 127 to residue 232 of SEQ. ID. NO. 301, residue 240 to residue 259 of SEQ. ID. NO. 301, residue 139 to residue 156 of SEQ. ID. NO. 295, residue 160 to residue 178 of SEQ. ID. NO. 295, residue 280 to residue 207 of SEQ. ID. NO. 295, residue 221 to residue 257 of SEQ. ID. NO. 295, residue 259 to residue 323 of SEQ. ID. NO. 295, residue 885 to residue 985 of SEQ. ID. NO. 299, residue 147 to residue 259 of SEQ. ID. NO. 363, residue 140 to residue 252 of SEQ. ID. NO. 344, residue 247 to

30

35

residue 356 of SEQ. ID. NO. 353, residue 359 to residue 391 of SEQ. ID. NO. 353, residue 120 to residue 254 of SEQ. ID. NO. 300, residue 287 to residue 311 of SEQ. ID. NO. 286, residue 313 to residue 352 of SEQ. ID. NO. 286, residue 354 to residue 401 of SEQ. ID. NO. 286, residue 208 to residue 252 of 5 SEQ. ID. NO. 287, residue 259 to residue 373 of SEQ. ID. NO. 287, residue 5 to residue 120 of SEQ. ID. NO. 293, residue 123 to residue 139 of SEQ. ID. NO. 293, residue 233 to residue 339 of SEQ. ID. NO. 265, residue 67 to residue 228 of SEQ. ID. NO. 278, residue 130 to residue 172 of SEQ. ID. NO. 274, residue 174 to residue 238 of SEQ. ID. NO. 274, residue 99 to residue 10 112 of SEQ. ID. NO. 274, residue 114 to residue 128 of SEQ. ID. NO. 274, residue 26 to residue 69 of SEQ. ID. NO. 285, residue 71 to residue 128 of SEQ. ID. NO. 285, residue 130 to residue 146 of SEQ. ID. NO. 285, residue 620 to residue 636 of SEQ. ID. NO. 327, residue 638 to residue 775 of SEQ. ID. NO. 327, residue 397 to residue 505 of SEQ. ID. NO. 301, residue 528 to residue 545 of SEQ. ID. NO. 301, residue 556 to residue 612 of SEQ. ID. NO. 15 301, residue 614 to residue 631 of SEQ. ID. NO. 301, residue 633 to residue 650 of SEQ. ID. NO. 301, residue 553 to residue 687 of SEQ. ID. NO. 299, residue 305 to residue 447 of SEQ. ID. NO. 289, residue 1 to residue 52 of SEQ. ID. NO. 364, residue 65 to residue 74 of SEQ. ID. NO. 364, residue 486 to residue 604 of SEQ. ID. NO. 275, residue 158 to residue 267 of SEQ. ID. 20 NO. 272, residue 270 to residue 282 of SEQ. ID. NO. 272, residue 163 to residue 237 of SEQ. ID. NO. 273, residue 240 to residue 251 of SEQ. ID. NO. 273, residue 213 to residue 344 of SEQ. ID. NO. 282, residue 183 to residue 324 of SEQ. ID. NO. 292, residue 327 to residue 341 of SEQ. ID. NO. 292, residue 352 to residue 372 of SEQ. ID. NO. 292, residue 141 to residue 166 of 25 SEQ. ID. NO. 271, residue 168 to residue 232 of SEQ. ID. NO. 271, residue 1 to residue 13 of SEQ. ID. NO. 302, residue 15 to residue 28 of SEQ. ID. NO. 302, residue 30 to residue 72 of SEQ. ID. NO. 302, residue 476 to residue 529 of SEQ. ID. NO. 277, residue 41 to residue 146 of SEQ. ID. NO. 299, residue 149 to residue 162 of SEQ. ID. NO. 299, residue 166 to residue 177 of SEQ. 30 ID. NO. 299, residue 192 to residue 203 of SEQ. ID. NO. 299, residue 71 to residue 343 of SEQ. ID. NO. 290, residue 346 to residue 363 of SEO. ID. NO. 290, residue 36 to residue 240 of SEQ. ID. NO. 331, residue 242 to residue 270 of SEQ. ID. NO. 331, residue 1 to residue 192 of SEQ. ID. NO. 375, residue 266 to residue 290 of SEQ. ID. NO. 375, residue 23 to residue 216 of 35 SEQ. ID. NO. 279, residue 220 to residue 270 of SEQ. ID. NO. 279, residue

285 to residue 386 of SEQ. ID. NO. 279, residue 84 to residue 234 of SEQ. ID. NO. 297, residue 248 to residue 259 of SEQ. ID. NO. 297, residue 261 to residue 269 of SEQ. ID. NO. 297, residue 275 to residue 402 of SEQ. ID. NO. 294, residue 1 to residue 171 of SEQ. ID. NO. 298, residue 403 to residue 417 of SEQ. ID. NO. 307, residue 420 to residue 453 of SEQ. ID. NO. 307, residue 456 to residue 464 of SEQ. ID. NO. 307, residue 468 to residue 690 of SEQ. ID. NO. 307, residue 1 to residue 285 of SEQ. ID. NO. 304, residue 287 to residue 315 of SEQ. ID. NO. 304, residue 318 to residue 336 of SEQ. ID. NO. 304, residue 255 to residue 269 of SEQ. ID. NO. 342, residue 271 to residue 10 337 of SEQ. ID. NO. 342, residue 347 to residue 467 of SEQ. ID. NO. 281, residue 116 to residue 136 of SEQ. ID. NO. 375, residue 138 to residue 357 of SEQ. ID. NO. 375, residue 133 to residue 423 of SEQ. ID. NO. 364, residue 141 to residue 299 of SEQ. ID. NO. 305, residue 202 to residue 365 of SEQ. ID. NO. 296, residue 134 to residue 426 of SEQ. ID. NO. 288, residue 1 to residue 218 of SEQ. ID. NO. 276, residue 1 to residue 246 of SEQ. ID. NO. **15** 280, residue 444 to residue 608 of SEQ. ID. NO. 364, residue 10 to residue 686 of SEQ. ID. NO. 283, residue 1 to residue 148 of SEQ. ID. NO. 296, residue 1 to residue 191 of SEQ. ID. NO. 287, residue 193 to residue 204 of SEQ. ID. NO. 287, residue 209 to residue 373 of SEQ. ID. NO. 287, residue 211 to residue 470 of SEQ. ID. NO. 284, residue 472 to residue 482 of SEQ. 20 ID. NO. 284, residue 133 to residue 144 of SEQ. ID. NO. 281, residue 146 to residue 336 of SEQ. ID. NO. 281, residue 1 to residue 264 of SEQ. ID. NO. 303, residue 265 to residue 295 of SEQ. ID. NO. 303, residue 297 to residue 326 of SEQ. ID. NO. 303, residue 328 to residue 338 of SEQ. ID. NO. 303, 25 residue 247 to residue 356 of SEQ. ID. NO. 353, residue 358 to residue 391 of SEQ. ID. NO. 353, residue 257 to residue 288 of SEQ. ID. NO. 298, residue 290 to residue 385 of SEQ. ID. NO. 298, residue 245 to residue 256 of SEQ. ID. NO. 298, residue 422 to residue 802 of SEQ. ID. NO. 303, residue 803 to residue 814 of SEQ. ID. NO. 303, residue 139 to residue 156 of SEQ. ID. NO. 295, residue 160 to residue 340 of SEQ. ID. NO. 295, residue 145 to residue 30 361 of SEQ. ID. NO. 282, residue 363 to residue 387 of SEQ. ID. NO. 282, residue 398 to residue 471 of SEQ. ID. NO. 282, residue 573 to residue 679 of SEQ. ID. NO. 320, residue 27 to residue 168 of SEQ. ID. NO. 291, residue 170 to residue 183 of SEQ. ID. NO. 291, residue 185 to residue 415 of SEQ. ID. 35 NO. 291, residue 1 to residue 301 of SEQ. ID. NO. 364, residue 114 to residue 702 of SEQ. ID. NO. 337, residue 377 to residue 412 of SEQ. ID. NO. 321,

10

15

20

25

30

35

residue 413 to residue 772 of SEQ. ID. NO. 321, residue 14 to residue 454 of SEQ. ID. NO. 265, residue 129 to residue 614 of SEQ. ID. NO. 268, residue 1 to residue 930 of SEQ. ID. NO. 300, residue 932 to residue 1046 of SEQ. ID. NO. 300, residue 1 to residue 301 of SEQ. ID. NO. 364, residue 1 to residue 42 of SEQ. ID. NO. 381, residue 44 to residue 973 of SEQ. ID. NO. 381, residue 1 to residue 93 of SEQ. ID. NO. 358, residue 179 of SEQ. ID. NO. 358, residue 179 of SEQ. ID. NO. 358, residue 181 to residue 227 of SEQ. ID. NO. 358, residue 114 to residue 702 of SEQ. ID. NO. 337, residue 1 to residue 659 of SEQ. ID. NO. 355, residue 661 to residue 907 of SEQ. ID. NO. 355, residue 1 to residue 131 of SEQ. ID. NO. 370, residue 133 to residue 601 of SEQ. ID. NO. 370, residue 1 to residue 412 of SEQ. ID. NO. 321, residue 413 to residue 772 of SEQ. ID. NO. 321, and residue 189 to residue 614 of SEQ. ID. NO. 364.

In a second aspect the present invention consists in a n isolated antigenic *Porphorymonas gingivalis* polypeptide, the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532 less the leader sequence set out in Table 3.

In a third aspect the present invention consists in an isolated DNA molecule, the DNA molecule comprising a nucleotide sequence which encodes the polypeptide of the first aspect the present invention or a sequence which hybridises thereto under stringent conditions.

It is preferred that the isolated DNA molecule comprises a nucleotide sequence selected from the group consisting of SEQ. ID. NO. 1 to SEQ. ID. NO. 264, SEQ. ID. NO. 529 and SEQ. ID. NO. 530.

In a fourth aspect the present invention consists in a recombinant expression vector comprising the DNA molecule of the second aspect of the present invention operably linked to a transcription regulatory element.

The present invention also provides a cell comprising this recombinant expression vector.

In a further aspect the present invention consists in a method for producing a *P. gingivalis* polypeptide comprising culturing the cell under conditions that permit expression of the polypeptide.

In yet a further aspect the present invention provides a composition for use in raising an immune response directed against *P. gingivalis* in a subject, the composition comprising an effective amount of at least one

10

15

20

25

polypeptide of the first aspect of the present invention, or at least one DNA molecule of the second aspect of the present invention, or both, and a pharmaceutically acceptable carrier. It is preferred that the pharmaceutically acceptable carrier is an adjuvant. In other aspects the present invention provides methods of treating *P. gingivalis* infection in subject comprising the administration of the composition to the subject such that treatment of *P. gingivalis* infection occurs. The treatment may be prophylactic or therapeutic.

In yet another aspect the present invention provides an antibody raised against a polypeptide of the first aspect the invention. The antibody may be polyclonal or monoclonal. The present invention also provides compositions including these antibodies. It is preferred that these compositions are adapted for oral use and may be, for example, dentrifices, mouthwashes, etc.

In a still further aspect the present invention provides a nucleotide probe comprising at least 18 nucleotides and having a contiguous sequence of at least 18 nucleotides identical to a contiguous nucleotide sequence selected from the group consisting of SEQ. ID. NO. 1 to SEQ. ID. NO. 121, SEQ. ID. NO. 529, and sequences complementary thereto. It is preferred that the probe further comprises a detectable label.

The present invention also provides a method for detecting the presence of *P. gingivalis* nucleic acid in a sample comprising:

- (a) contacting a sample with the nucleotide probe under conditions in which a hybrid can form between the probe and a *P. gingivalis* nucleic acid in the sample; and
- (b) detecting the hybrid formed in step (a), wherein detection of a hybrid indicates the presence of a *P. gingivalis* nucleic acid in the sample.

30 DETAILED DESCRIPTION

Definitions

A purified or isolated polypeptide or a substantially pure preparation of a polypeptide are used interchangeably herein and, as used herein, mean a polypeptide that has been separated from other proteins, lipids, and nucleic acids with which it naturally occurs. Preferably, the polypeptide is also

35

10

15

20

25

30

35

separated from substances, e.g., antibodies or gel matrix, e.g., polyacrylamide, which are used to purify it. Preferably, the polypeptide constitutes at least 10, 20, 50 70, 80 or 95% dry weight of the purified preparation. Preferably, the preparation contains: sufficient polypeptide to allow protein sequencing; at least 1, 10, or 100 mg of the polypeptide.

A purified preparation of cells refers to, in the case of plant or animal cells, an in vitro preparation of cells and not an entire intact plant or animal. In the case of cultured cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

A purified or isolated or a substantially pure nucleic acid, e.g., a substantially pure DNA, (are terms used interchangeably herein) is a nucleic acid which is one or both of the following: not immediately contiguous with both of the coding sequences with which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the naturally occurring genome of the organism from which the nucleic acid is derived; or which is substantially free of a nucleic acid with which it occurs in the organism from which the nucleic acid is derived. The term includes, for example, a recombinant DNA which is incorporated into a vector, e.g., into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other DNA sequences. Substantially pure DNA also includes a recombinant DNA which is part of a hybrid gene encoding additional *P. gingivalis* DNA sequence.

A "contig" as used herein is a nucleic acid representing a continuous stretch of genomic sequence of an organism.

An "open reading frame", also referred to herein as ORF, is a region of nucleic acid which encodes a polypeptide. This region may represent a portion of a coding sequence or a total sequence and can be determined from a stop to stop codon or from a start to stop codon.

As used herein, a "coding sequence" is a nucleic acid which is transcribed into messenger RNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the five prime terminus and a translation stop code at the three prime terminus. A

10

15

20

25

30

35

coding sequence can include but is not limited to messenger RNA synthetic DNA, and recombinant nucleic acid sequences.

A "complement" of a nucleic acid as used herein refers to an antiparallel or antisense sequence that participates in Watson-Crick base-pairing with the original sequence.

A "gene product" is a protein or structural RNA which is specifically encoded by a gene.

As used herein, the term "probe" refers to a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest. Probes are often associated with or capable of associating with a label. A label is a chemical moiety capable of detection. Typical labels comprise dyes, radioisotopes, luminescent and chemiluminescent moieties, fluorophores, enzymes, precipitating agents, amplification sequences, and the like. Similarly, a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest and immobilizes such molecule is referred herein as a "capture ligand". Capture ligands are typically associated with or capable of associating with a support such as nitro-cellulose, glass, nylon membranes, beads, particles and the like. The specificity of hybridization is dependent on conditions such as the base pair composition of the nucleotides, and the temperature and salt concentration of the reaction. These conditions are readily discernible to one of ordinary skill in the art using routine experimentation.

Homologous refers to the sequence similarity or sequence identity between two polypeptides or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two sequences divided by the number of positions compared x 100.

The terms peptides, proteins, and polypeptides are used interchangeably herein.

An "immunogenic component" as used herein is a moiety, such as an *P. gingivalis* polypeptide, analog or fragment thereof, that is capable of eliciting a humoral and/or cellular immune response in a host animal.

10

15

20

25

30

35

An "antigenic component" as used herein is a moiety, such as *P. gingivalis* polypeptide, analog or fragment thereof, that is capable of binding to a specific antibody with sufficiently high affinity to form a detectable antigen-antibody complex.

As used herein, the term "cell-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

As used herein, the term "control sequence" refers to a nucleic acid having a base sequence which is recognized by the host organism to effect the expression of encoded sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include a promoter, ribosomal binding site, terminators, and in some cases operators; in eukaryotes, generally such control sequences include promoters, terminators and in some instances, enhancers. The term control sequence is intended to include at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences.

As used herein, the term "operably linked" refers to sequences joined or ligated to function in their intended manner. For example, a control sequence is operably linked to coding sequence by ligation in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequence and host cell.

A "sample" as used herein refers to a biological sample, such as, for example, tissue or fluid isolated from an individual (including without limitation plasma. serum, cerebrospinal fluid, lymph, tears, saliva and tissue sections) or from *in vitro* cell culture constituents, as well as samples from the environment.

The practice of the invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, A Practical Guide to Molecular

15

20

25

30

35

Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, Volumes 1 and 2, IRL Press (1991), D.M. Glover and B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996), and F.M. Ausubel et al. (Editors), Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience (1988, including all updates until present). The disclosure of these texts are incorporated herein by reference.

10 Pharmaceutically Acceptable Carriers

The antibodies, polypeptides and DNA of the present invention can be included in compositions which include a carrier or diluent. These compositions include pharmaceutical compositions where the carrier or diluent will be pharmaceutically acceptable. Pharmaceutically acceptable carriers or diluents include those used in compositions suitable for oral, rectal, nasal, topical (including buccal and sublingual), vaginal, parenteral (including subcutaneous, intramuscular, intravenous, intradermal, intrathecal and epidural) administration. They are non-toxic to recipients at the dosages and concentrations employed. Representative examples of pharmaceutically acceptable carriers or diluents include, but are not limited to; water, isotonic solutions which are preferably buffered at a physiological pH (such as phosphate-buffered saline or Tris-buffered saline) and can also contain one or more of, mannitol, lactose, trehalose, dextrose, glycerol, ethanol or polypeptides (such as human serum albumin). The compositions may conveniently be presented in unit dosage form and may be prepared by any of the methods well known in the art of pharmacy.

As will be well understood by those skilled in the art alterations may be made to the amino acid sequences set out in the Sequence Listings. These alterations may be deletions, insertions, or substitutions of amino acid residues. The altered polypeptides can be either naturally occurring (that is to say, purified or isolated from a natural source) or synthetic (for example, by performing site-directed metagenesis on the encoding DNA). It is intended that such altered polypeptides which have at least 85%, preferably at least 95% identity with the sequences set out in the Sequence Listing are

10

20

25

30

35

within the scope of the present invention. Antibodies raised against these altered polypeptides will also bind to the polypeptides having one of the sequences set out in the Sequence Listings. The level of % identity is to be calculated as set out above.

Protein sequences are homologous if they are related by divergence from a common ancestor. Consequently, a species homologue of the protein will be the equivalent protein which occurs naturally in another species. Within any one species a homologue may exist as numerous allelic variants, and these will be considered homologues of the protein. Allelic variants and species homologues can be obtained by following standard techniques known to those skilled in the art.

An allelic variant will be a variant that is naturally occurring within an individual organism.

15 Mutants, Variants and Homology - Nucleic Acids

Mutant polynucleotides will possess one or more mutations which are deletions, insertions, or substitutions of nucleotide residues. Mutants can be either naturally occurring (that is to say, isolated from a natural source) or synthetic (for example, by performing site-directed metagenesis on the DNA). It is thus apparent that polynucleotides of the invention can be either naturally occurring or recombinant (that is to say prepared using recombinant DNA techniques).

An allelic variant will be a variant that is naturally occurring within an individual organism.

Nucleotide sequences are homologous if they are related by divergence from a common ancestor. Consequently, a species homologue of the polynucleotide will be the equivalent polynucleotide which occurs naturally in another species. Within any one species a homologue may exist as numerous allelic variants, and these will be considered homologues of the polynucleotide. Allelic variants and species homologues can be obtained by following standard techniques known to those skilled in the art.

Antibody Production

Antibodies, either polyclonal or monoclonal, which are specific for a polypeptide of the present invention can be produced by a person skilled in the art using standard techniques such as, but not limited to, those described

10

15

20

25

30

35

by Harlow et al. Antibodies: A Laboratory Manual, Cold Springs Harbor Laboratory Press (1988), and D. Catty (editor), Antibodies: A Practical Approach, IRL Press (1988).

Various procedures known in the art may be used for the production of polyclonal antibodies to epitopes of a protein. For the production of polyclonal antibodies, a number of host animals are acceptable for the generation of antibodies by immunization with one or more injections of a polypeptide preparation, including but not limited to rabbits, mice, rats, etc. Various adjuvants may be used to increase the immunological response in the host animal, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminium hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, oil emulsions, keyhole lympet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

A monoclonal antibody to an epitope of a protein may be prepared by using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Kohler and Milstein (1975, Nature 256, 493-497), and the more recent human B-cell hybridoma technique (Kesber et al. 1983, Immunology Today 4:72) and EBVhybridoma technique (Cole et al. 1985, Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. pp. 77-96). In addition, techniques developed for the production of "chimeric antibodies" by splicing the genes from antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity may be used (Morrison et al. 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger et al. 1984 Nature 312:604-608; Takeda et al. 1985 Nature 31:452-454). Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce 4-specific single chain antibodies.

Recombinant human or humanized versions of monoclonal antibodies are a preferred embodiment for human therapeutic applications. Humanized antibodies may be prepared according to procedures in the literature (e.g. Jones et al. 1986, Nature 321:522-25; Reichman et al. 1988 Nature 332:323-27; Verhoeven et al. 1988, Science 239:1534-36). The

10

15

25

30

35

recently described "gene conversion metagenesis" strategy for the production of humanized monoclonal antibody may also be employed in the production of humanized antibodies (Carter et al. 1992 Proc. Natl. Acad. Sci. U.S.A. 89:4285-89). Alternatively, techniques for generating the recombinant phase library of random combinations of heavy and light regions may be used to prepare recombinant antibodies (e.g. Huse et al. 1989 Science 246:1275-81).

Antibody fragments which contain the idiotype of the molecule such as Fu F(ab1) and F(ab2) may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab) E2 fragment which can be produced by pepsin digestion of the intact antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragment, and the two Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent. Alternatively, Fab expression libraries may be constructed (Huse et al. 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragment with the desired specificity to a protein.

Adjuvants

20 "Adjuvant" means a composition comprised of one or more substances that enhances the immunogenicity and efficacy of a vaccine composition. Non-limiting examples of suitable adjuvants include squalane and squalene (or other oils of animal origin); block copolymers; detergents such as Tween®-80; Quil® A, mineral oils such as Drakeol or Marcol, vegetable oils such as peanut oil; Corynebacterium-derived adjuvants such as Corynebacterium parvum; Propionibacterium-derived adjuvants such as Propionibacterium acne; Mycobacterium bovis (Bacillus Calmetic and Guerinn or BCG); interleukins such as interleukin 2 and interleukin-12; monokines such as interleukin 1; tumour necrosis factor; interferons such as gamma interferon; combinations such as saponin-aluminium hydroxide or Quil-A aluminium hydroxide; liposomes; ISCOM adjuvant; mycobacterial cell wall extract; synthetic glycopeptides such as muramyl dipeptides or other derivatives; Avridine; Lipid A; dextran sulfate; DEAE-Dextran or DHAE-Dextran with aluminium phosphate; carboxypolymethylene such as Carbopol' EMA; acrylic copolymer emulsions such as Neocryl A640 (e.g. U.S.

Pat. No. 5,047,238); vaccinia or animal posvirus proteins; sub-viral particle adjuvants such as cholera toxin, or mixtures thereof.

As used herein, stringent conditions are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% NaDodSO4 at 50° C; (2) employ during hybridisation a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS and 10% dextran sulfate at 42°C in 0.2 x SSC and 0.1% SDS

15

10

5

As will be understood the present invention includes within its scope DNA vaccination. Further information regarding DNA vaccination may be found in Donnelly et al, Journal of Immunological Methods 176(1994) 145-152, the disclosure of which is incorporated herein by reference.

20

25

30

35

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer, or group of elements or integers.

Preparation of the P. gingivalis library for sequencing.

To determine the DNA sequence of *P. gingivalis* genomic DNA was isolated from *P. gingivalis* strain W50 (ATCC 53978) essentially by the method described by Mamur J. (J. Mol. Biol. 3, 208-218, 1961). Cloning of DNA fragments was performed essentially as described by Fleischmann *et al.*, (Science; 269, 496-512, 1995)(2). Briefly, purified genomic DNA from *P. gingivalis* was nebulized to fragment the DNA and was treated with Bal31 nuclease to create blunt ends then run twice through preparative 1% agarose gels. DNA fragments of 1.6-2.0 kb were excised from the gel and the DNA recovered. This DNA was then ligated to the vector pUC18 (*Sma*I digested and dephosphorylated; Pharmacia) and electrophoresed through a 1%

preparative agarose gel. The fragment comprising linear vector plus one insert was excised, purified and this process repeated to reduce any vector without insert contamination. The recovered vector plus insert DNA was blunt-ended with T4 DNA polymerase, then a final ligation to produce circular DNA was performed. Aliquots of Epicurian Coli Electroporation-Competent Cells (Stratagene) were transformed with the ligated DNA and plated out on SOB agar antibiotic diffusion plates containing X-gal and incubated at 37°C overnight. Colonies with inserts appeared white and those without inserts (vector alone) appeared blue. Plates were stored at 4°C until the white clones were picked and expanded for the extraction of plasmid DNA for sequencing.

DNA sequencing

10

15

20

25

30

Plasmid DNA was prepared by picking bacterial colonies into 1.5ml of LB, TB or SOB broth supplemented with 50-100ug/ml Ampicillin in 96 deep well plates. Plasmid DNA was isolated using the QIAprep Spin or QIAprep 96 Turbo miniprep kits (QIAGEN GmbH, Germany). DNA was eluted into a 96 well gridded array and stored at -20C.

Sequencing reactions were performed using ABI PRISM Dye
Terminator and ABI PRISM BIGDye Terminator Cycle Sequencing Ready
Reaction kits with AmpliTaq DNA polymerase FS (PE Applied Biosystems,
Foster City, CA) using the M13 Universal forward and reverse sequencing
primers. Sequence reactions were conducted on either a Perkin-Elmer
GeneAmp 9700 (PE Applied Biosystems) or Hybaid PCR Express (Hybaid,
UK) thermal cyclers. Sequencing reactions were analysed on ABI PRISM 377
DNA sequencers (PE Applied Biosystems).

The sequences obtained are set out below. The relationship between these sequences is set out in Table 1. The initiation codon was calculated using a combination of sequence homology alignment (FASTA), signal sequence prediction (PSORT, SignalP) or ORF prediction (GeneMark).

Table 1: Reference table indicating the relationships of each sequence ID to the selected proteins.

| Protein name | DNA sequence of | Amino acid sequence of | DNA sequence of protein | Amino acid sequence of |
|--------------|-----------------|------------------------|-------------------------|------------------------|
| | complete ORF | complete ORF | - | protein |
| PG1 | 1 | 265 | 122 | 386 |
| PG10 | 2 | 266 | 123 | 387 |
| PG100 | 3 | 267 | 124 | 388 |
| PG101 | 4 | 268 | | |
| PG102 | 5 | 269 | 125, 126 | 389, 390 |
| PG104 | 6 | 270 | 127 | 391 |
| PG105 | 7 | 271 | 128 | 392 |
| PG106 | 8 | 272 | 129 | 393 |
| PG107 | 9 | 273 | 130, 131, 132 | 394, 395, 396 |
| PG108 | 10 | 274 | 133 | 397 |
| PG109 | 11 | 275 | 134, 135 | 398, 399 |
| PG11 | 12 | 276 | 136 | 400 |
| PG110 | 13 | 277 | 137 | 401 |
| PG111 | 14 | 278 | | |
| PG112 | 15 | 279 | 138, 139 | 402, 403 |
| PG113 | 16 | 280 | 140 | 404 |
| PG114 | 17 | 281 | 141 | 405 |
| PG115 | 18 | 282 | 142 | 406 |
| PG116 | 19 | 283 | 143 | 407 |
| PG117 | 20 | 284 | 144 | 408 |
| PG118 | 21 | 285 | 145 | 409 |
| PG119 | 22 | 286 | 146 | 410 |
| PG12 | 23 | 287 | 147 | 411 |
| PG120 | 24 | 288 | 148 | 412 |

| Protein | DNA | Amino acid | DNA sequence of | Amino acid |
|---------|-------------|--------------|---------------------|-------------|
| name | sequence of | sequence of | protein | sequence of |
| | complete | complete ORF | _ | protein |
| | ORF | | | |
| PG121 | 25 | 289 | 149 | 413 |
| PG122 | 26 | 290 | 150 | 414 |
| PG123 | 27 | 291 | 151 | 415 |
| PG124 | 28 | 292 | 152 | 416 |
| PG125 | 29 | 293 | 153 | 417 |
| PG126 | 30 | 294 | 154 | 418 |
| PG13 | 31 | 295 | 155 | 419 |
| PG14 | 32 | 296 | 156 | 420 |
| PG15 | 33 | 297 | 1 5 <i>7</i> | 421 |
| PG16 | 34 | 298 | 158 | 422 |
| PG18 | 35 | 299 | 159 | 423 |
| PG2 | 36 | 300 | 160, 161 | 424, 425 |
| PG21 | 37 | 301 | 162 | 426 |
| PG22 | 38 | 302 | 163 | 427 |
| PG23 | 39 | 303 | 164 | 428 |
| PG24 | 40 | 304 | 165 | 429 |
| PG25 | 41 | 305 | 166 | 430 |
| PG27 | 42 | 306 | 167 | 431 |
| PG28 | 43 | 307 | 168 | 432 |
| PG29 | 44 | 308 | 169 | 433 |
| PG3 | 45 | 309 | 170 | 434 |
| PG30 | 46 | 310 | 171 | 435 |
| PG31 | 47 | 311 | 172 | 436 |
| PG32 | 48 | 312 | 173 | 437 |
| PG33 | 49 | 313 | 174 | 438 |
| PG34 | 50 | 314 | 175, 176 | 439, 440 |
| PG35 | 51 | 315 | 177 | 441 |
| PG36 | 52 | 316 · | 178 | 442 |
| PG37 | 53 | 317 | 179, 180 | 443, 444 |
| PG38 | 54 | 318 | 181 | 4 45 |

| Protein name | DNA sequence of | | DNA sequence of protein | Amino acid sequence of |
|-----------------|-----------------|--------------|-------------------------|------------------------|
| | complete | complete ORF | P. 0.0.2 | protein |
| | ORF | 1 | | • |
| PG39 | 55 | 319 | 182 | 446 |
| PG4 | 56 | 320 | 183 | 447 |
| PG40 | 5 <i>7</i> | 321 | 184 | 448 |
| PG41 | 58 | 322 | 185 | 449 |
| PG42 | 59 | 323 | 186 | 450 |
| PG43 | 60 | 324 | 187 | 451 |
| PG44 | 61 | 325 | 188 | 452 |
| PG45 | 62 | 326 | 189 | 453 |
| PG46 | 63 | 327 | 190 | 454 |
| PG47 | 64 | 328 | 191 | 455 |
| PG48 | 65 | 329 | 192 | 456 |
| PG49 | 66 | 330 | 193 | 4 5 <i>7</i> |
| PG5 | 67 | 331 | 194 | 458 |
| PG50 | 68 | 332 | 195 | 459 |
| PG51 | 69 | 333 | 196 | 460 |
| PG52 | 70 | 334 | 197 | 461 |
| PG53 | 71 | 335 | 198 | 462 |
| PG54 | 72 | 336 | 199 | 463 |
| PG55 | 73 | 337 | 200 | 464 |
| PG56 | 74 | 338 | 201, 202 | 465, 466 |
| PG57 | 75 | 339 | 203, 204, 205 | 467, 468, 469 |
| PG58 | 76 | 340 | 206, 207 | 470, 471 |
| PG59 | 77 | 341 | 208, 209, 210 | 472, 473, 474 |
| PG6 | 78 | 342 | 211 | 475 |
| PG60 | 79 | 343 | 212 | 476 |
| PG61 | 80 | 344 | 213 | 477 |
| PG62 | 81 | 345 | 214 | 478 |
| PG63 | 82 | 346 | 215 | 479 |
| PG64 | 83 | 347 | 216 | 480 |
| PG65 | 84 | 348 | 217 | 481 |

| Protein | DNA | Amino acid | DNA sequence of | Amino acid |
|---------|-------------|--------------|-----------------|--------------------|
| name | sequence of | | protein | sequence of |
| | complete | complete ORF | 1 - | protein |
| | ORF | • | | r |
| PG66 | 85 | 349 | 218 | 482 |
| PG67 | 86 | 350 | 219 | 483 |
| PG68 | 87 | 351 | 220, 221 | 484, 485 |
| PG69 | 88 | 3 52 | 222 | 486 |
| PG7 | 89 | 353 | 223 | 487 |
| PG70 | 90 | 354 | 224 | 488 |
| PG71 | 91 | 355 | 225 | 489 |
| PG72 | 92 | 356 | 226 | 490 |
| PG73 | 93 | 357 | 227 | 491 |
| PG74 | 94 | 358 | 228 | 492 |
| PG75 | 95 | 359 | 229 | 493 |
| PG76 | 96 | 360 | 230 | 494 |
| PG77 | 97 | 361 | 231 | 495 |
| PG78 | 98 | 362 | 232 | 496 |
| PG79 | 99 | 363 | 233 | 497 |
| PG8 | 100 | 364 | 234, 235, 236, | 498, 499, 500, 501 |
| | | | 237 | |
| PG80 | 101 | 365 | 238 | 50 2 |
| PG81 | 102 | 366 | 102 | 366 |
| PG82 | 103 | 367 | 239 | 503 |
| PG83 | 104 | 368 | 240 | 504 |
| PG84 | 105 | 369 | 241, 242 | 505, 506 |
| PG85 | 106 | 370 | 243 | 507 |
| PG86 | 107 | 371 | 244, 245 | 508, 509 |
| PG87 | 108 | 372 | 246 | 510 |
| PG88 | 109 | 373 | 247, 248, 249 | 511, 512, 513 |
| PG89 | 110 | 374 | 250 | 514 |
| PG9 | 111 | 375 | 251, 252, 253 | 515, 516, 517 |
| PG90 | 112 | 376 | 254, 255 | 518, 519 |
| PG91 | 113 | 377 | 256 | 520 |

| Protein name | DNA sequence of complete ORF | ì | DNA sequence of protein | Amino acid sequence of protein |
|-----------------|------------------------------|-----|-------------------------|--------------------------------------|
| PG92 | 114 | 378 | 257 | 521 |
| PG93 | 115 | 379 | 258 | 522 |
| PG94 | 116 | 380 | 259 | 523 |
| PG95 | 117 | 381 | 260 | 524 |
| PG96 | 118 | 382 | 261 | 525 |
| PG97 | 119 | 383 | 262 | 526 |
| PG98 | 120 | 384 | 263 | 527 |
| PG99 | 121 | 385 | 264 | 528 |
| PG127 | 529 | 531 | 530 | 532 |

DNA sequence analysis

5

10

DNA files in FASTA format were converted to GCG format files and imported into a database. The DNA files were translated into amino acid files using the program Flip obtained from ANGIS(Australian Genomic Information Service, University of Sydney, Australia). A series of bioinformatic analyses were performed on the proteins in order to select potential vaccine candidates. The programs used were FASTA homology searching (1), PSORT (2,3), SignalP (4), TopPred (5), and GeneMark (6). The proteins and their bioinformatic results were stored in the custom written database for search and retrieval of proteins with the desired characteristics

15

20

The FASTA homology results for these proteins were then examined for any alignment with a protein suggesting surface location or vaccine efficacy. All proteins were searched for homology against a non-redundant bacterial protein database compiled by ANGIS using the FASTA algorithm. The settings used for the FASTA searches were Ktup = 2, gap creation penalty = -12, gap extension penalty = -2, width for deriving alignment in opt = 16 and the Blosum 50 scoring matrix. Individual FASTA search results were examined for significant homology by statistical probability and amino acid alignments. The results are set out in Table 2.

10

15

20

25

30

Protein files were then trimmed to the first, second, third, fourth and fifth methionine residues using a protein trimming program (ANGIS). The trimmed proteins were then subjected to PSORT analysis for the detection of signal sequences and the prediction of cell location. Proteins exhibiting a PSORT probability of outer membrane >0.8 were considered to indicate surface localisation. A second signal sequence detection program SignalP was also performed and, in certain instances, this program detected signals not identified with PSORT. All proteins identified by other methods were also analysed by PSORT and SignalP. Previously, the C-terminal amino acid of bacterial outer membrane proteins has been shown to be important for the assembly of the protein on the outer membrane (7). A typical structure definition for outer membrane proteins has been determined as the presence of a signal sequence at the N-terminus and a tyrosine or phenylalanine at the C-terminus. A number of the selected proteins exhibit this characteristic structure. The program TopPred was used to determine the presence and number of membrane spanning domains (MSDs) and the presence of such sequences indicates a preference to be attached to membranes such as the outer membrane. The results of PSORT, SignalP and TopPred analyses with the C-terminal amino acids of the selected proteins are set out in Table 3.

The 70 amino acids from the C-terminus of a number of *P. gingivalis* outer membrane proteins share 50-100% protein sequence identity. These proteins included RGP1, RGP2, KGP, HagA, HagC, HagD, prtH and prtT. This conserved motif may be involved in the attachment or sorting of proteins to the outer membrane. The protein data set was searched using FASTA homology as described above and a number of novel proteins were identified which demonstrate similar motifs at their C-termini. The results are listed in Table 4

The TonBIII box is a 30 amino acid motif present within TonB outer membrane receptors in a wide variety of bacteria. The TonBIII box of *P. gingivalis* (8) was used to search the protein data set for homology by FASTA as described above. Those proteins demonstrating significant homology are listed in Table 5.

Table 2: FASTA protein homology results of complete ORFs against a non-redundant protein database.

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | ology resi | ılts |
|---------|---|---------------------|-----------|--------------------------------|----------------------------|---------------|-----------|
| паше | | accession number | homolog | accession homolog gingivalis | | | |
| | | | | | Identity % Overlap E value | Overlap | E value |
| PG1 | 48kD outer membrane protein, Actinobacillus pleuropneumoniae | U24492 | 449aa | 451aa | 32 | 4 54aa | 1.40E-42 |
| PG2 | Outer membrane protein (susC), Bacteroides thetaiotaomicron | L49338 | 1038aa | 1017aa, 1014aa | 28 | 1099aa | 4.60E-32 |
| PG3 | Outer membrane porin F adhesin, Pseudomonas fluorescens | U19743 | 317aa | 223aa | 35 | 187aa | 1.10E-10 |
| PG4 | Outer membrane protein A. Escherichia fergusonii | M63352 | 243aa | 672aa | 48 | 88aa | 4.10E-10 |
| PG5 | Adhesin protein (AdcA), Streptococcus pneumoniae | 271552 | 423aa | 315aa | 25 | 279aa | 9.40E-15 |
| PG6 | Hemolysin A (phyA). Prevotella melaninogenica | U27587 | 332аа | 324aa | 09 | 306aa | 3.00E-74 |
| PG7 | Hemolysin (tlyC), Serpulina hyodysenteriae | X73141 | 268aa | 404aa | 33 | 266aa | 1.40E-24 |
| PG8 | Heme uptake protein A, Bacteriodes fragilis | X97122 | 431aa | 598aa, 550aa, 458aa, 426aa | 79 | 417aa | 6.70E-121 |
| PG9 | PG9 Internalin A (inlA), Lysteria monocytogenes | M67471 | 744aa | a, 1232aa, a | 38 | 340aa | 7.30E-23 |

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | ology resu | ılts |
|---------|--|-------------------|-----------|--------------------------------|---|------------|-----------|
| name | | accession homolog | | gingivalis | | | |
| | | number | | protein | | | |
| | | | | | Identity % Overlap E value | Overlap | E value |
| PG10 | Macrophage infectivity potentiator (MIP). Legionella oakridgensis. | U92214 | 234aa | 195aa | 50 | 201aa | 4.70E-31 |
| PG11 | Haemagglutinin (phg), Prevotella intermedia | AF017417 | 309aa | 313aa | 44 | 309aa | 3.60E-44 |
| PG12 | Outer membrane lipoprotein, Haemophilus | M68502 | 274aa | 271aa | 36 | 254aa | 9.60E-27 |
| | influenzae | | | | | | |
| PG13 | Ferric receptor (cfr.A), Campylobacter coli | U80812 | 696аа | 757aa | 24 | 625aa | 1.20E-18 |
| PG14 | 36kD antigen. Helicobacter pylori | U86610 | 329aa | 331aa | 37 | 326aa | 1.10E-35 |
| PG15 | Outer membrane protein. Erwinia amylovara | X77921 | 377aa | 267aa | 30 | 253aa | 5.40E-08 |
| PG16 | C terminal protease, Bartonella bacilliformis | L37094 | 434aa | 569aa | 36 | 357aa | 3.00E-35 |
| PG18 | Protein-export membrane protein (secD), Helicobacter AE000652 | | 503aa | 981aa | 32 | 611aa | 1.10E-36 |
| | pylori | | | | *************************************** | | |
| PG21 | Surface antigen gene, Methanosarcina mazei | X84710 | 783aa | 821aa | 37 | 331aa | 6.20E-33 |
| PG22 | Alpha-hemolysin gene, Aeromonas hydrophila | L36462 | 85aa | 106aa | 57 | 67aa | 2.60E-14 |
| PG23 | clpA/clpB protease. Bacillus subtilis | D26185 | 810aa | в59аа | 45 | 855aa | 7.10E-122 |
| PG24 | Putative hemolysin, Streptococcus mutans | AF051356 | 445aa | 417aa | 29 | 432aa | 1.80E-29 |

WO 99/29870

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | ology resu | ılts |
|---------|--|---------------------|---|---|------------------------|-------------|----------|
| name | | accession number | homolog | accession homolog gingivalis number protein | : | ; ; ; | |
| | | | | | Identity % Overlap | | E value |
| PG25 | Cysteine protease, Porphyromonas gingivalis | U54691 | 1723aa | 293aa | 42 | 142aa | 1.10E-12 |
| PG27 | TonB linked adhesin. Porphyromonas gingivalis | Y07618 | 1097aa | 312aa | 45 | 360aa | 3.20E-41 |
| PG28 | Cysteine protease / hemagglutinin, Porphyromonas | S75942 | 886aa | 843aa | 35 | 838aa | 7.00E-90 |
| PG30 | Putative NIpD lipoprotein, Aquifex aeolicus | AE000754 | 187aa | 337aa | 42 | 142aa | 1.80E-12 |
| PG31 | | X73141 | 141aa | | 31 | 123aa | 1.80E-07 |
| PG32 | lomonas | • | 350aa | 391aa | 26 | 382aa | 3.40E-07 |
| | aeruginosa | | | | | | |
| PG33 | Major outer membrane protein (oprF),Pseudomonas fluorescens | L21200 | 317aa | 385aa | 32 | 163aa | 2.30E-06 |
| PG34 | mbrane protein, Rhodobacter capsulatus | Q07396 | 193aa | 190aa | 46 | 190aa | 2.20E-36 |
| PG35 | Colcin 1 receptor, Escherichia coli | J04229 | 663aa | 833aa | 25 | 590aa | 2.40E-10 |
| PG36 | Outer membrane antigen (oma87), Pasteurella | U60439 | 789aa | 891aa | 21 | 894aa | 3.70E-10 |
| | multocida | | *************************************** | | | | |
| PG37 | Cationic outer membrane protein (ompH), Yersinia | M34854 | 164aa | 174aa, 170aa | 27 | 168aa | 4.30E-07 |
| | enterocolitica | | | | | | |

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | ology resu | ılts |
|---------|---|-----------|-------------------|--------------------------------|------------------------|------------|----------|
| name | | accession | accession homolog | gingivalis | | } | |
| | | number | | protein | | | |
| | | | | | Identity % Overlap | | E value |
| PG38 | Cationic outer membrane protein (ompH), Yersinia enterocolitica | M34854 | 164aa | 163aa | 23 | 160aa | 5.90E-05 |
| PG39 | Outer membrane protein (susC) Bacteroides thetaiotaomicron | L49338 | 1038aa | 827aa | 24 | 347aa | 1.50E-06 |
| PG40 | Heme receptor (Hut A), Vibrio cholera | Q56644 | 693aa | 772aa | 23 | 722aa | 4.90E-09 |
| PG41 | Outer membrane protein (tolC), Escherichia coli | X54049 | 495aa | 462aa | 22 | | 4.60E-09 |
| PG42 | Neuraminidase, Micromonospora viridifaciens | D01045 | 647aa | 492aa | 32 | 375aa | 2.10E-22 |
| PG43 | Immunoreactive outer membrane protein (omp28), Brucella melitensis | U30815 | 250aa | 245aa | 24 | | 0.0015 |
| PG44 | Macrophage infectivity potentiator, Legionella israelensis | U92208 | 242aa | 276aa | 35 | 219aa | 9.10E-18 |
| PG45 | Outer membrane protein, Neisseria meningitidis | AF021245 | 797aa | 775aa | 21 | 699aa | 0.0034 |
| PG46 | Outer membrane protein 85, Neisseria gonorrhoeae | UB1959 | 792aa | 774aa | | | 0.00098 |
| PG47 | Outer membrane protein (susC) Bacteroides | L49338 | 1038aa | 867aa | | | 1.00E-03 |
| | IIICIOIIICIOII | | | | | | |

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | nology resu | ılts |
|---------|---|-------------------|-----------|--------------------------------|---|-------------|-----------|
| name | | accession homolog | | gingivalis | | | |
| | | numper | | protein | | | |
| | | | | | Identity % Overlap E value | Overlap | E value |
| PG48 | Immunoglobulin binding surface protein (sir22), | X75750 | 365aa | 431aa | 25 | 269aa | 5.20E-05 |
| | Streptococcus pyogenes | | | | *************************************** | | |
| PG49 | Fimbrillin (orf2), Porphyromonas gingivalis | D42067 | 453aa | 333aa | 23 | 296aa | 0.062 |
| PG50 | Outer membrane protein (susC) Bacteroides | L49338 | 1038aa | 848aa | 26 | 579aa | 1.60E-11 |
| | thetaiotaomicron | | | | | | |
| PG51 | PGaA antigen. Porphyromonas gingivalis | X95938 | 202aa | 202aa | 54 | 126aa | 1.20E-25 |
| PG52 | Alkaline protease secretion apparatus (aprF) | X64558 | 481aa | 455aa | 21 | 427aa | 3.50E-06 |
| | Pseudomonas aeruginosa | | | | | | |
| PG53 | Protein export protein (tolC), Salmonella enteritidis | U25178 | 491aa | 444aa | 23 | 436aa | 6.20E-11 |
| PG54 | Protease I. Achromobacter lyticus | J5128 | 653aa | 940aa | 24 | 695aa | 1.50E-22 |
| PG55 | Fimbrillin (orf3), Porphyromonas gingivalis | D42067 | 670aa | 670aa | 43 | б88аа | 4.90E-106 |
| PG56 | Cysteine protease Porphyromonas gingivalis | U68468 | 364aa | 1282aa, 1274aa | 25 | 212aa | 0.00012 |
| PG57 | Cysteine protease, Porphyromonas gingivalis | U68468 | 1358aa | 924aa, 922aa, 921aa | 31 | 742aa | 1.40E-23 |
| PG60 | Outer membrane protein 11, Helicobacter pylori | AE000562 | 186aa | 547aa | 25 | 183aa | 2.20E+00 |
| PG61 | Ferric pseudobactin M114 receptor protein (pbuA), | X73412 | 826aa | 749aa | 22 | 585aa | 1.00E-05 |
| | Pseudomonas sp. | | | | | | |

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | ology resu | ılts |
|---------|--|-----------|-------------------|--------------------------------|----------------------------|------------|----------|
| name | | accession | accession homolog | gingivalis | | | |
| | | number | | protein | | | |
| | | | | | Identity % Overlap E value | Overlap | E value |
| PG66 | Attachment and invasion protein (ail), Salmonella typhimurium | AF007380 | 165aa | 206aa | 21 | 140aa | 1.90E+00 |
| PG68 | factor, Streptococcus pyogenes | U02290 | 1025aa | 1225aa. 1224aa | 24 | 176aa | 2.10E-01 |
| PG69 | Vacuolating cytotoxin (vacA), Helicobacter pylori | U63261 | 160aa | 425aa | 32 | 111aa | 1.20E+00 |
| PG70 | Outer membrane protein, Neisseria gonorrhoea | U52069 | 174aa | 266aa | 22 | 153aa | 6.90E+00 |
| PG71 | Gliding motility protein (gldA), Flavobacterium iohnsoniae | AF007381 | 578aa | 834aa | 23 | 572aa | 3.90E-25 |
| PG75 | Class 3 outer membrane porin (porB), Neisseria meningitidis | U07191 | 332aa | 391aa | 23 | 239aa | 4.60E-01 |
| PG81 | ane protein (ompA),Shigella dysenteriae | V01344 | 351aa | >235aa | 26 | 186aa | 3.10E-01 |
| PG82 | Outer membrane protein (alkL), Pseudomonas oleovorans | X65936 | 230aa | 434aa | 56 | 136aa | 2.80E+00 |
| PG83 | Gliding motility protein (gldA), Flavobacterium johnsoniae | AF007381 | 578aa | 926aa | 21 | 639aa | 8.50E-09 |
| PG87 | rium tuberculosis | AL021942 | 877aa | 781aa | 29 | 794aa | 2.20E-34 |

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | ology resu | ılts |
|---------|--|-------------------|-----------|--------------------------------|------------------------|------------|----------|
| name | | accession homolog | | gingivalis nrotoin | | | |
| | | | | | Identity % | Overlap | E value |
| PG89 | NADH-ubiquinone oxidoreductase, Helicobacter pvlori | AE000631 | 512aa | 259aa | 24 | 186aa | 3.90E-01 |
| PG91 | Neuraminidase (nanH), Bacteroides fragilis | D28493 | 544aa | 540aa | 24 | 251aa | 1.60E+00 |
| PG92 | Hypothetical protein, Mycobacterium tuberculosis | AL021942 | 877aa | 771aa | 29 | 770aa | 8.00E-30 |
| PG93 | ae | X07191 | 219aa | 776aa | 41 | 63аа | 6.90E-01 |
| PG94 | ······································ | D26470 | 991aa | 1157aa | 24 | - | 7.60E-08 |
| PG95 | Sensor protein (EVGS). Escherichia coli | D14008 | 1197aa | 961aa | 28 | 511aa | 2.60E-17 |
| PG105 | Plasma cell membrane glycoprotein, Human | P22413 | 873aa | 449aa | 34 | 404aa | 5.60E-33 |
| PG106 | Hypothetical secreted protein. Helicobacter pylori | O24951 | 242aa | 246aa | 30 | 252aa | 7.80E-22 |
| PG107 | otein, Mycobacterium | 032883 | 229aa | 246aa, 241aa, 232aa | 46 | 193aa | 1.20E-26 |
| | ap Idai | | | | | | |
| PG108 | ABC transporter, Archaeoglobus fulgidis | 029244 | 228aa | 219aa | 51 | 219aa | 3.80E-41 |
| PG109 | Proteinase IV, Escherichia coli | F64936 | 618aa | 595aa, 589aa | 38 | 597aa | 1.10E-57 |
| PG110 | Preprotein translocase, Staphylococcus aureus | 006446 | 843aa | 523aa | 43 | 521aa | 6.00E-71 |
| PG111 | ABC transporter, Synechocystis sp. | P73758 | 574aa | >720aa | 40 | 579aa | 1.70E-73 |
| PG112 | Glycosyl transferase, Erwinia anylovora | Q46634 | 351aa | 375aa, 362aa | 31 | | 1.60E-32 |

| Protein | Protein Homology description | Genbank | Length of | Genbank Length of Length of P. | FASTA homology results | ology resu | ılts |
|---------|---|-----------|-------------------|--------------------------------|----------------------------|------------|-----------|
| name | | accession | accession homolog | gingivalis | | | |
| | | number | | protein | | | |
| | | | | | Identity % Overlap E value | Overlap | E value |
| PG113 | Heat shock protein (dnak), Treponema pallidum | AE001203 | 635aa | 640aa | 62 | 644aa | 9.10E-138 |
| PG114 | Dihydrolipamide dehydrogenase, Clostridium | Q59299 | 578aa | 449aa | 37 | 450aa | 3.80E-54 |
| | magnum | | | | | | |
| PG115 | Zinc protease, Escherichia coli | P31828 | 931aa | 941aa | 27 | 890aa | 6.60E-57 |
| PG116 | Heat shock protein (HTPG), Escherichia coli | P10413 | 624aa | 684aa | 32 | 627aa | 4.60E-48 |
| PG117 | Transcriptional regulator, Aquifex aeolicus | 066591 | 506aa | 464aa | 39 | 389aa | 2.40E-49 |
| PG118 | ABC transporter, Bacillus subtilus | H70019 | 261aa | 250aa | 59 | 251aa | 1.50E-60 |
| PG119 | ATP-dependent protease, Aquifex aeolicus | 066827 | 444aa | 461aa | 46 | 458aa | 1.60E-77 |
| PG120 | Nitrogen assimilation regulatory protein. | P10576 | 480aa | 457aa | 49 | 242aa | 3.80E-45 |
| | Bradvrhizobium sp. | | | | | | |
| PG121 | Cobalamin synthesis protein, Bacillus megaterium | E1331323 | 367aa | 602aa | 36 | 324aa | 9.20E-37 |
| PG122 | Outer membrane integrity (tolA), Haemophilus | P71397 | 819aa | 443aa | 37 | 441aa | 1.90E-54 |
| | influenzae | | | | | | |
| PG123 | Fimbrillin. Porphyromonas gingivalis | D1034032 | 490aa | 479aa | 32 | 480aa | 7.30E-48 |
| PG124 | Heat shock protein (dnaf). Leptospira interrogans | AF007813 | 369aa | 383aa | 46 | | 2.30E-57 |

| name PG125 Cobalami | | accession | |) - | * | 3 | |
|---------------------|--|-----------|---------|------------------------------|----------------------------|---------|-----------|
| | | | homolog | accession homolog gingivalis | | | |
| | | number | | protein | | | |
| | | | | | Identity % Overlap E value | Overlap | E value |
| | | Q05592 | 264aa | 293aa | 37 | 259aa | 3.70E-26 |
| typhimur | typhimurium | | | | | | |
| PG126 ABC-type | ABC-type permease, Pseudomonas aeruginosa | 068878 | 326aa | 356aa | 33 | 333aa | 1.30E-30 |
| PG127 Endonucl | Endonuclease excision repair protein (uvrB), | X93486 | 670aa | 678aa | 56 | 675aa | 1.10E-134 |
| Pseudoing | Pseudomonas aeruginosa | | | | | | |

SUBSTITUTE SHEET (Rule 26) (RO/AU)

signal sequence as determined by PSORT. The cell location & probability values are generated by PSORT and represent space (PC) or cytoplasm (C). The number of transmembrane domains (TMDs) was determined by TopPred and does not the probability of the protein being in the cell compartments outer membrane (OM), inner membrane (IM), periplasmic presence of a signal sequence detected with either PSORT or SignalP. The terms in parentheses indicates the type of Table 3: Results of PSORT, SignalP and TopPred analysis of the proteins. The signal present column indicates the include uncleavable signal sequences.

| al Number | Amino Acid of TMD's | | | 0 | က | | 8 | 0 | 0 | 1 | 8 | 0 | c |
|-----------------------------|---------------------|--------|----|-------|--------|--------|-----------------|-----------------|-------|-------|-------|-------|----------|
| C-terminal | Amino Ac | | | Z. | [14 | ĹĿ. | × | Δ. | æ | S | ĿЛ | z | 2 |
| | | | C | 0.22 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0 | 0.22 | 0.25 |
| bability | | | PS | 0 | 0.33 | 0.29 | 0 | 0. | 0 | 0 | 0 | 0. | 0 |
| Cell Location & probability | | | IM | 0 | 0 | 0 | 0.76 | 0.7 | 0.25 | 0 | 0.42 | 0 | 0 |
| Cell Loca | | | ОМ | 0 | 0.94 | 0.94 | 0.79 | 0.79 | 0 | 0 | 0 | 0 | 0 |
| PSORT | cleavage | site | | 34 | 20 | 17 | 18 | 22 | 35 | | | | 1 |
| SignalP | cleavage | site | | 24 | 20 | 17 | | 22 | 40 | 1 | 7 | | <u>.</u> |
| Methionine | in ORF | | | 1 | 1 | 2 | 11 | 1 | 1 | 1 | 1 | 1 | 2 |
| Signal Present | | | | Ą | γ | Ϋ́ | Y (lipoprotein) | Y (lipoprotein) | γ | z | z | z | z |
| Protein | Length | | | 451aa | 1017аа | 1014aa | 223aa | 672aa | 315aa | 324aa | 404ав | 598aa | 550aa |
| | Glpas | number | | 386 | 424 | 425 | 434 | 447 | 458 | 475 | 487 | 498 | 499 |
| Protein Protein | пате | | | PG1 | PG2 | PG2 | PG3 | PG4 | PGs | PG6 | PG7 | PG8 | PGs |

| Protein Protein | ł | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Loca | Cell Location & probability | bability | | C-terminal | Number |
|-----------------|--------|---------|-----------------|------------|----------|----------|-----------|-----------------------------|----------|------|------------|----------|
| name | SeqID | Length | | in ORF | cleavage | cleavage | | | | | Amino Acid | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | _ | OM | IM | PS | ပ | | |
| PG8 | 500 | 458aa | Z | 3 | | , | 0 | 0 | 0 | 0.34 | z | 0 |
| PG8 | 501 | 426aa | Z | 4 | | | 0 | o | 0 | | 1 | 0 |
| | 515 | 1266aa | Z | 1 | 7 | | 0 | | 0 | 0.22 | | 1 |
| | 516 | 1232aa | Z | 2 | | | 0 | o | | 0.39 | Ĺλ | Ę. |
| PG9 | | 1174aa | Z | 3 | | | 0 | 0 | 0 | 0.47 | ធា | 1 |
| 0 | | 195aa | Z | 1 | , | | | | 0 | 0.11 | × | 0 |
| | 400 | 313aa | Y | - | 22 | 26 | _ | | 0.93 | 0 | Ж | 1 |
| PG12 | 411 | 271aa | Y (lipoprotein) | 3 | 27 | 29 | | 0.7 | 0 | 0 | 跃 | 0 |
| PG13 | 419 | 757aa | γ | 1 | 23 | 25 | | 0 | 0.29 | 0 | z | 0 |
| PG14 | 420 | 331aa | Y (uncleavable) | - | 35 | 26 | | 0.58 | 0 | 0 | × | 1 |
| PG15 | 421 | 267aa | Ą | 2 | 24 | 18 | 0 | | 0 | 0 | × | 1 |
| PG16 | 422 | 569aa | Y (lipoprotein) | 1 | 24 | 18 | _ | | 0 | 0 | ڻ | 0 |
| PG18 | 423 | 981aa | γ | | 30 | | 0 | 0.56 | 0 | 0 | × | 11 |
| PG21 | 426 | 821aa | Ą | 2 | 24 | 27 | | 0 | 0.37 | o | | 1 |
| PG22 | 427 | 106аа | Y (uncleavable) | - | | 41 | | 0.29 | 0 | 0 | | 0 |
| PG23 | 428 | 859aa | Z | - | | | 0 | 0.12 | o | 0 | Ą | -1 |

| Protein Protein | Protein | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Local | Cell Location & probability | bability | | C-terminal | Number |
|-----------------|----------------|---------|-----------------|-------------|----------|----------|------------|-----------------------------|----------|------|---------------|----------|
| name | SeqID | Length | | in ORF | cleavage | cleavage | - | | | | Amino Acid | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | | OM | M | PS | U | | |
| PG24 | 429 | 417aa | Y | | 19 | 19 | 0 | 0.44 | 0 | 0 | z | 3 |
| PG25 | 430 | 293aa | Y | 1 | 27 | | 0.2 | 0 | 0.62 | 0 | ĸ | 0 |
| PG27 | 431 | 312аа | Z . | - | | | | 0 | 0 | 0.28 | O | 1 |
| PG28 | 432 | 843aa | Y | e 1 | 21 | 21 | 0.93 | 0 | 0.24 | 0 | I | 1 |
| PG29 | 133 | 290aa | Y | 1 | 18 | | 0.28 | 0 | 0.94 | 0 | Х | - |
| PG30 | 435 | 337aa | Y | - | 21 | 21 | 0.24 | 0 | | 0 | Х | 0 |
| PG31 | 436 | 15 1aa | z | | | , | 0 | 0 | | 0.3 | H | 0 |
| PG32 | 437 | 391aa | Y | 1 | 20 | 20 | 0.62 | 0 | 0.13 | 0 | У | 0 |
| PG33 | 438 | 385aa | γ | | 26 | 56 | | 0 | 0.31 | 0 | ъ | 1 |
| PG34 | 439 | 190aa | Y | | , | 13 | 0 | 0.5 | 0 | 0 | ¥ | വ |
| PC34 | 440 | 186ва | Y (uncleavable) | 2 | | | 0 | 0.5 | 0 | 0 | Y | 4 |
| PG35 | 441 | 833aa | Υ | 1 | 22 | 22 | 0.94 | 0 | 0.37 | 0 | г. | -1 |
| PG36 | 442 | 891aa | Y (uncleavable) | , - | , | 40 | 0 | 0.31 | 0 | 0 | Ŀ, | 2 |
| PG37 | PG37 443 174aa | | Y (uncleavable) | | 28 | 24 | | 0.35 | 0 | 0 | × | 0 |

| Protein Protein | 1 | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Locat | Cell Location & probability | ability | | C-terminal | Number |
|-----------------|--------|---------|-----------------|--------------|----------|----------|------------|-----------------------------|---------|------|------------|-------------|
| name | SeqID | Length | | in ORF | cleavage | cleavage | | | | | Amino Acid | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | _ | МО | IM | PS | IJ | | |
| PG37 | 444 | 170aa | Y (uncleavable) | 2 | 24 | 20 | 0 | 0.35 | 0 | 0 | ~ | 0 |
| | | | Y | _ | 18 | | 0.21 | 0 | 0.93 | 0 | Y | 1 |
| PG39 | 446 | 827aa | Y | ~ | 36 | 36 | 0.93 | 0 | 0.25 | 0 | Į, | 3 |
| | | 772ва | Ϋ́ | 2 | | | | 0 | 0.32 | 0 | ഥ | 4 |
| | | 462aa | Å | 2 | | | 0.25 | 0 | 0.54 | 0 | δ | 2 |
| | | 492aa | Å | 5 | 30 | | | 0 | 0.00 | 0.13 | | 2 |
| | | | Y (uncleavable) | 2 | | 22 | | ~ | | 0 | X | 1 |
| PG44 | | 276аа | Y | - | 19 | | | 0 | | 0 | Ж | 0 |
| | | | Y (lipoprotein) | - | | | | 0.7 | o, | 0 | £2 | 4 |
| | | | Y | 1 | | | | o | 0.22 | 0 | ഥ | 2 |
| PG47 | 455 | 867aa | γ | -1 | 24 | 24 | | 0 | 0.38 | 0 | Į. | 2 |
| | 456 | 431aa | γ | 1 | 24 | 24 | 0 | 0.1 | 0 | 0 | ᄶ | |
| PG49 | 457 | 333аа | Y (uncleavable) | | 24 | 18 | | 0.12 | 0 | 0 | | 0 |
| PG50 | 459 | 848aa | Y | - | 21 | 21 | 0.94 | 0 | 0.34 | 0 | 4 | 6 |
| PG51 | 460 | 202ав | γ | 1 | 26 | 25 | 0.2 | 0 | 0.61 | 0 | S | 0 |
| PG52 | 461 | 455aa | Y (uncleavable) | - | 23 | 21 | 0 | 0.18 | 0 | 0 | 뵤 | 1 |

| Protein Protein | Protein | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Loca | Cell Location & probability | bability | ; | C-terminal | Number |
|-----------------|---------|---------|-----------------|------------|----------|----------|-----------|-----------------------------|----------|-------|--------------|----------|
| name | seqID | Length | | in ORF | cleavage | cleavage | | | | | Amino Acid | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | | МО | IM | PS | S | | |
| PG53 | 462 | 444aa | γ | | 14 | 17 | 0.36 | 0 | 0.22 | 0 | Ω | 2 |
| PG54 | 463 | 940aa | Y | 1 | 27 | 20 | 0.86 | 0 | 0.25 | 0 | ŏ | 5 |
| PG55 | 464 | 670aa | Y (lipoprotein) | _ | 23 | 23 | 0.79 | 0.7 | 0 | 0 | 7 | 2 |
| PG56 | 465 | 1282aa | Y (uncleavable) | 1 | | 21 | 0 | 0.04 | 0 | 0 | Х | 4 |
| PG56 | 466 | 1274aa | Z | 2 | | | 0 | 0 | 0 | 0.27 | × | ιn. |
| | 467 | 925aa | Å | | 28 | 24 | 0.53 | 0 | 0.2 | 0 | d | က |
| | | 922aa | Ą | 2 | | 21 | 0.53 | 0 | 0.2 | 0 | Ь | က |
| PG57 | 469 | 921aa | Y | 3 | 24 | 20 | 0.53 | 0 | 0.2 | 0 | Д | က |
| | 470 | 593aa | Y | 1 | 24 | 24 | 0.82 | 0 | 0.19 | 0 | ഥ | 1 |
| PG58 | 471 | 589aa | Y | 2 | 20 | 20 | 0.82 | 0 | 0.19 | 0 | ഥ | -1 |
| PG59 | 472 | 346aa | Y | | 37 | 1 | 0 | 0.18 | 0 | 0 | 17- | 1 |
| PG59 | 473 | 345aa | Y | 2 | 36 | 56 | 0.92 | 0 | 0.15 | 0 | ഥ | 1 |
| PG59 | 474 | 330aa | Ÿ | 3 | 21 | 41 | 0.93 | 0 | 0.25 | 0 | Ţ. | 1 |
| PG60 | 476 | 547aa | γ | 1 | 28 | 28 | 0.93 | 0 | 0.25 | 0 | E. | 0 |
| PG61 | 477 | 749aa | Y | 2 | 21 | 21 | 0.94 | 0 | 0.29 | 0 | ഥ | 3 |
| PG62 | 478 | 494aa | Y | 1 | 21 | 21 | 0.93 | 0 | 0.24 | 0 | [<u>t</u> . | 2 |

| Protein Protein | Protein | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Local | Cell Location & probability | sability | | C-terminal | Number |
|-----------------|---------|---------|----------------|------------------|----------|----------|------------|-----------------------------|----------|---|-------------|----------|
| пате | SeqID | Length | | in ORF | cleavage | cleavage | | | | | Amino Acid | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | | МО | IM | PS | C | | |
| PG63 | 479 | 294aa | Y Y | 1 | 20 | 20 | 0.93 | 0 | 0.24 | 0 | Œ, | 1 |
| PC64 | 480 | 204aa | y | 1 | 20 | 20 | 0.93 | 0 | 0.19 | 0 | ഥ | 1 |
| PG65 | 481 | 243aa | Y | 1 | 18 | | 0.93 | 0 | 0.25 | 0 | ţr. | 1 |
| PG66 | 482 | 206aa | Y | 1 | 21 | 21 | 0.94 | 0 | 0.3 | 0 | ĿĻ | 1 |
| PG67 | 483 | 950aa | Y | 1 | 28 | 36 | | 0 | 0.27 | 0 | Y | 4 |
| | | 1226aa | Y | 1 | | | | 0 | 0.31 | 0 | Ϋ́ | 0 |
| | | 1225aa | Y | 2 | | | 0.91 | 0 | 0.31 | 0 | Y | 0 |
| PG69 | 486 | 425aa | Y | 1 | | 29 | | 0 | 0.21 | 0 | <u>tr</u> , | 7 |
| PG70 | 488 | 260aa | Y | - | | 24 | | 0 | 0.24 | 0 | ĹĽ | 0 |
| PG71 | 489 | 834aa | Y | 2 | 20 | 20 | | 0 | 0.31 | 0 | z | 2 |
| PG72 | 490 | 399aa | γ | 1 | | 27 | | 0 | 0.32 | 0 | H | 2 |
| PG73 | 491 | 382aa | γ | 2 | | 20 | | 0 | 0.3 | 0 | L) | 1 |
| PG74 | 492 | 222аа | γ | 1 | 24 | 24 | | 0 | 0.32 | 0 | _l | 0 |
| PG75 | 493 | 391aa | γ | - | 26 | 26 | | 0 | 0.3 | 0 | I | 1 |
| PG76 | 494 | 446aa | Å | , - - | | 22 | | 0 | 0.32 | 0 | Λ | က |
| PG77 | 495 | 308аа | γ | 2 | | 28 | | | 0.38 | 0 | × | 0 |





of TMD's Amino Acid **C-terminal** O' 0.12 C Cell Location & probability 0.25 0.25 0.46 0.03 0.21 0.21 0.21 0.25 0.23 0.25 0.2 PS \mathbf{X} MO 0.94 0.93 0.93 0.93 0.93 0.93 0.93 0.93 0.93 0.91 cleavage PSORT site 20 24 57 22 23 20 46 47 22 cleavage SignalP 16 20 44 23 22 23 24 Methionine in ORF Signal Present Protein 240aa 926aa 400ав 398aa 239аа 211aa 270aa 781aa 267ав Protein number SeqID 497 366 503 504 505 506 508 509 514 507 Protein name PG79 PG80 PC83 PG81 PG82 PG84 PG84 PG85 PG86 PG86 PG87 PG88 PG88 PG88



| Protein Protein | i | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Loca | Cell Location & probability | bability | | C-terminal | Number |
|-----------------|--------|---------|--------------------|------------------|----------|----------|-----------|-----------------------------|----------|---|------------|----------|
| name | seqID | Length | | in ORF | cleavage | cleavage | | | | | Amino Acid | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | | МО | IM | PS | C | | |
| PG90 | 518 | 229aa | Y | 1 | 22 | 21 | 0.85 | 0 | 0.44 | 0 | × | 0 |
| PG90 | 519 | 228aa | Ą | 2 | 21 | | 0.85 | 0 | 0.44 | 0 | У | 0 |
| PG91 | 520 | 540aa | Å | | | | 0.85 | 0 | 0.30 | 0 | щ | 0 |
| PG92 | 521 | 771aa | Y | 2 | | | 0.85 | 0 | 0.3 | 0 | ж | က |
| PG93 | 522 | 776aa | Y | 1 | | | 0.85 | 0 | 0.37 | 0 | R | 4 |
| PG94 | 523 | 1157aa | Ą | | 23 | 28 | 0.8 | 0 | 0.25 | 0 | ð | 5 |
| PG95 | 524 | 961aa | Y (lipoprotein) | , - 1 | | | 0.79 | 0.87 | 0 | 0 | Λ | 1 |
| PG96 | 525 | 563аа | Å | Ε. | 23 | | | 0 | 0.33 | 0 | × | 0 |
| PG97 | 526 | 437аа | Ą | - | | | | 0 | 0.65 | 0 | ď | 0 |
| PG98 | 527 | 318aa | Y (lipoprotein) | | 19 | 19 | 0.79 | 0.7 | 0 | 0 | L. | - |
| PG99 | 528 | 461aa | Y (uncleavable) | 1 | | 20 | 0 | 0 | 0.3 | 0 | R | 0 |
| PG100 | 388 | 279аа | γ | 1 | 20 | 18 | 0.26 | 0 | 0.54 | 0 | jament. | 0 |
| PC101 | 268 | >157aa | N (ORF incomplete) | | | | | | | | × | 1 |
| PC102 | 389 | 562aa | Ϋ́ | 1 | 29 | 29 | 0.19 | 0 | 0.4 | 0 | S | 3 |
| PG102 | 390 | 558aa | Y | 2 | 25 | 25 | | 0 | 0.46 | 0 | S | က |
| PG104 | 391 | 391aa | Ą | | | | 0.62 | 0 | 0.22 | | × | 0 |

| Protein Protein | Protein | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Loca | Cell Location & probability | bability | | C-terminal | Number |
|-----------------|---------|---------|--------------------|-------------|----------|----------|-----------|-----------------------------|----------|------|------------|----------------|
| пате | QIpas | Length | - | in ORF | cleavage | cleavage | | | | | Amino Acid | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | | МО | M | PS | ပ | | |
| PG105 | 392 | 449aa | Y | | 22 | 19 | 0.31 | 0 | 0.91 | 0 | Ъ | က |
| PG106 | 393 | 246aa | Å | 2 | 41 | | 0 | 0 | 0 | 0.02 | П | 0 |
| PG107 | 394 | 246aa | Z | 1 | , | | 0 | 0 | 0 | 0.32 | D | - |
| PG107 | 395 | 241aa | Z | 2 | , | | 0 | 0 | 0 | 0.3 | Ω | 1 |
| PG107 | 396 | 232aa | z | 3 | | | 0 | 0 | 0 | 0.21 | Ω | |
| PC108 | 397 | 219aa | Z | 1 | , | | 0 | 0 | 0 | 0.19 | R | , , |
| PG109 | 398 | 595aa | Y | 1 | 35 | 37 | 9 | 0 | 0.93 | 0 | Y | က |
| PG109 | 399 | 589аа | Ϋ́ | 2 | 29 | 31 | | 0 | 0.93 | 0 | Y | 3 |
| PC110 | 401 | >523aa | z | 1 | 1 | | 0 | 0 | 0 | 0.38 | incomplete | 0 |
| PG111 | 278 | >720aa | N (ORF incomplete) | • | , | ı | | | | | U | H |
| PC112 | 402 | 375aa | γ | 1 | | 43 | 0 | 0.12 | 0 | 0 | z | -1 |
| PG112 | 403 | 362aa | ¥ | 7 | | 30 | 0 | 0 | 0.12 | 0 | z | 1 |
| PC113 | 404 | 640aa | z | 1 | | | 0 | | 0 | 0.25 | × | 1 |
| PG114 | 405 | 449aa | z | 1 | | | 0 | | 0 | 0 | ပ | 4 |
| PC115 | 406 | 941aa | Y | 1 | 23 | 22 | 0.13 | 0 | 0.92 | 0 | ď | 2 |
| PG116 | 407 | 684aa | Z | 9 -4 | | | | 0.12 | 0 | 0 | L | 2 |

| Protein | Protein Protein | Protein | Signal Present | Methionine | SignalP | PSORT | Cell Local | Cell Location & probability | ability | | C-terminal | Number |
|---------|-----------------|---------|----------------|------------|----------|----------|------------|-----------------------------|---------|------|---------------------|----------|
| name | SeqID | Length | | in ORF | cleavage | cleavage | | | | | Amino Acid of TMD's | of TMD's |
| | number | | | | site | site | | | | | | |
| | | | | | | | ОМ | IM | PS | C | | |
| PG117 | 408 | 464aa | Z | 1 | | | 0 | 0.19 | 0 | 0 | -1 | F-4 |
| PC118 | 409 | 250aa | Z | 1 | | | 0 | 0 | 0 | 0.27 | ш | |
| PG119 | 410 | 461aa | Z | _ | | 1 | 0 | 0.28 | 0 | 0 | ET) | 2 |
| PG120 | 412 | 457aa | Z | | | | 0 | | 0 | 0.21 | ш | 0 |
| PG121 | 413 | 602aa | Z | <u>—</u> | | | 0 | 0 | 0 | 0.31 | ш | က |
| PG122 | 414 | 443aa | Z | — | | 1 | 0 | 0 | 0 | 0.14 | O | বা |
| PG123 | 415 | 479aa | Y | 2 | 22 | 22 | 0.26 | 0 | 0.94 | 0 | × | 0 |
| PG124 | 416 | 383aa | Z | 1 | , | | 0 | 0 | 0 | 0.29 | D | 2 |
| PG125 | 417 | 293aa | Ą | 1 | 23 | 15 | 0.18 | 0 | 0.93 | 0 | ~ | 1 |
| PG126 | 418 | З5баа | z | | | | 0 | 0.52 | 0 | 0 | D | 6 |
| PG127 | 532 | 678aa | 678aa N | — | ! | | 0 | o | 0 | 0.28 | V | 2 |

Table 4: Percentage identity and percentage similarity of various proteins with the 70 amino acids from the C-terminal of the *P. gingivalis* arginine protease 1 (RGP1), arginine protease 2 (RGP2), and the cysteine protease/hemagglutinin (prtT).

| ٠ | - |
|---|---|
| | |
| | |

| Protein name | P | ercent ide | ntity | Pe | rcent sim | ilarity |
|-----------------|------|------------|-------|------|-----------|------------|
| | RGP1 | RGP2 | prtT | RGP1 | RGP2 | prtT |
| PG21 | 17 | 29 | 21 | 40 | 57 | 4 9 |
| PG25 | 43 | 41 | 9 | 64 | 73 | 14 |
| PG27 | 41 | 33 | 7 | 73 | 74 | 11 |
| PG28 | 21 | 26 | 34 | 49 | 57 | 74 |
| PG54 | 19 | 13 | 16 | 40 | 43 | 33 |
| PG57 | 11 | 14 | 19 | 20 | 24 | 34 |
| PG91 | 31 | 21 | 39 | 57 | 53 | 74 |
| PG96 | 0 | 13 | 20 | 0 | 24 | 43 |
| PG97 | 10 | 26 | 33 | 14 | 47 | 61 |
| PG98 | 16 | 20 | 0 | 47 | 54 | 0 |
| PG99 | 19 | 0 | 26 | 41 | 0 | 54 |
| PG100 | 20 | 21 | 24 | 39 | 57 | 41 |
| PG101 | 11 | 16 | 27 | 17 | 39 | 60 |
| PG102 | 27 | 20 | 31 | 50 | 61 | 61 |
| PG104 | 16 | 23 | 26 | 46 | 44 | 49 |

Table 5: Percentage identity and percentage similarity of various proteins with the TonBIII box of *P. gingivalis*.

| Protein name | Percent identity | Percent similarity |
|-----------------|------------------|--------------------|
| PG2 | 46 | 71 |
| PG13 | 57 | 93 |
| PG35 | 50 | 96 |
| PG47 | 39 | 71 |
| PG50 | 54 | 93 |

Cloning, expression and purification of recombinant P. gingivalis genes.

PG₁

Oligonucleotides to the 5' and 3' regions of the deduced protein were 5 used to amplify the gene of interest from a preparation of P. gingivalis W50 genomic DNA using the TaqPlus Precision PCR System (Stratagene) and a PTC-100 (M) Research) thermal cycler or similar device. The 5' oligonucleotide primer sequence was GCGCCATATGCTGGCCGAACCGGCC, the 3' oligonucleotide primer sequence was 10 GCGCCTCGAGTCAATTCATTTCCTTATAGAG. The PCR fragment was purified, digested with Nde I, Xho I restriction enzymes (Promega) and ligated into the corresponding sites of the plasmid pProEx-1 (Gibco-BRL) and transformed into E. coli ER1793 cells (a gift from Elizabeth Raleigh, New England Biolabs). A resulting clone expressing the correct insert was 15 selected and induced with or without 0.1mM IPTG (Promega) for expression of the recombinant protein. Expression of the recombinant protein was determined by SDS-PAGE analysis and Western Blot using the one of the rabbit antisera described above or an anti-hexahistidine antibody (Clontech) 20 that detects the hexahistidine tag that was fused to the P. gingivalis recombinant protein. PG1 was purified by disruption of the E. coli cells by sonication in binding buffer (Novagen) and solubilisation by the addition of sarkosyl (N-Lauroyl sarcosine) to a 1% final concentration. There after the preparation was diluted to 0.1% sarkosyl in binding buffer, bound to a 25 Nickel-nitrilotriacetic acid column (Ni-NTA; Qiagen), after washing bound proteins were eluted with 1M imidazole in elution buffer (Novagen) according to the Qiagen recommendations with 0.1% sarkosyl added to all buffers. Following purification samples were dialysed against 500mM NaCl, 20mM Tris, 0.1% sarkosyl at pH7.4 to remove the imidazole, concentrated as required and stored at 4°C until used. Purity and antigenicity were assessed 30 by SDS-PAGE and Western blot using selected antisera (from those described above) and the protein concentration was determined by the BCA assay (Pierce).

45

PG₂

5

10

20

25

30

35

The methods used for PG2 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACATGAAAAGAATGACGC, the 3' oligonucleotide primer sequence was CGCGAGATCTGAAAGACAACTGAATACC and the PCR product was cloned into pGex-stop RBS(IV) (Patent application WO9619496, JC Cox, SE Edwards, I Frazer and EA Webb. Variants of human papilloma virus antigens) using the BstZ 171 and Bgl II restriction sites. 2% sarkosyl was used to solubilise PG2 and 8M urea was added to the solublisation buffer and to all other buffers. Urea was removed from the purified protein by sequential dialysis (4M then 2M then 1M then 0.5M then 0M urea all in 50mM Tris, 500mM NaCl, 0.1% sarkosyl, pH7.4). Purified protein was stored at 4°C until required.

15 **PG3**

The methods used for PG3 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGTATACATGAAGAAATCAAGTGTAG, the 3' oligonucleotide primer sequence was GCGCAGATCTCTTCAGCGTACCTTGCTGTG and DNA was amplified with Pfu DNA polymerase (Stratagene). The PCR product was cloned directly into pCR-Blunt and transformed into E. coli Top10F'(InVitrogen) before subcloning into the expression plasmid pGexstop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into E. coli BL21DE3 (Pharmacia Biotech). The following modifications were made to the purification of PG3 from the PG1 method. Cells expressing the recombinant protein were disrupted by sonication in binding buffer and the insoluble inclusion bodies concentrated by centrifugation. Inclusion bodies were then solubilised in 6M urea (Sigma) in binding buffer and eluted with 6M urea added to the elution buffer. In some instances 6M guanidine hydrochloride (Sigma) was used instead of urea for these steps. Urea (or guanidine hydrochloride when it was substituted) was removed from the purified protein by sequential dialysis against reducing levels of urea (3M then 1.5M then 0.5M then 0M urea all in 50mM Tris, 500mM NaCl, 8% glycerol, pH7.4). Purified protein was stored frozen at -80°C until required. Protein concentration was determined by the Coomassie Plus protein assay (Pierce).

BNSDOCID: <WO___9929870A1_I_>

PG4

The methods used for PG4 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CTTCTGTATACTTACAGCGGACATCATAAAATC, the 3' oligonucleotide primer sequence was TTCCAGGAGGGTACCACGCAACTCTTCTTCGAT and DNA was amplified with the Tth XL PCR kit (Perkin Elmer). The PCR product was cloned into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Kpn I restriction sites and transformed into *E. coli* ER1793.

10 **PG**5

15

20

25

The methods used for PG5 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was TTGCAACATATGATCAGAACGATACTTTCA, the 3' oligonucleotide primer sequence was AGCAATCTCGAGCGGTTCATGAGCCAAAGC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24 (Novagen) using the Nde I and Xho I restriction sites and transformed into *E. coli* BL21 (Pharmacia Biotech). Removal of urea was not proceeded past 1M urea as the protein was insoluble at lower concentrations of urea. Purified protein was stored at 4°C until required.

PG6

The methods used for PG6 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was TAAACATATGTGCCTCGAACCCATAATTGCTCCG, the 3' oligonucleotide primer sequence was CGTCCGCGGAAGCTTTGATCGGCCATTGCTACT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Nde I and Hind III restriction sites and transformed into *E. coli* BL21.

30 PG8

The methods used for PG8 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACATGGAGTTCAAGATTGTG, the 3' oligonucleotide primer sequence was CGCGAGATCTGTTTTCTGAAAGCTTTTC and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was

cloned into the expression plasmid pProEx-1 using the Nde I and Xho I restriction sites and transformed into *E. coli* ER1793.

PG8A

5

10

PG8A is a shortened version of PG8 and has the first 173 amino acids removed. The methods used for PG8A were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACATGGAAAACTTAAAGAAC, the 3' oligonucleotide primer sequence was CGCGAGATCTGTTTTCTGAAAGCTTTTC and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was cloned into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* ER1793. Prior to dialysis of the purified protein EDTA (Sigma) was added to a final concentration of 10mM.

15

PG10

The methods used for PG10 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGATATCATGGATAAAGTGAGCTATGC, the 3' oligonucleotide primer sequence was CGCGAGATCTTTTGTTGATACTCAATAATTC and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was digested with Eco RV and Bgl II and ligated into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* ER1793.

25

20

PG11

The methods used for PG11 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGTATACATGAGAGCAAACATTTGGCAGATACTTTCCG, the 3' oligonucleotide primer sequence was GCGCAGATCTGCGCAAGCGCAGTATATCGCC and DNA was amplified with Tli DNA polymerase (Promega). The PCR product was cloned into pCR-Blunt and transformed into E. coli Top10F'before subcloning into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into E. coli ER1793. PG11 was purified by solubilisation of E. coli cells with 2% sarkosyl in binding buffer (Qiagen) which was diluted to

0.1% sarkosyl in binding buffer, bound to a Nickel-nitrilotriacetic acid column (Ni-NTA; Qiagen), after washing bound proteins were eluted with 1M imidazole (0.7% CHAPS (Sigma) in elution buffer; Qiagen) according to the Qiagen recommendations. Following purification samples were dialysed against 500mM NaCl, 20mM Tris, 0.7% CHAPS, 20% glycerol (Sigma) at pH7.4 to remove the imidazole, concentrated as required and stored at 4°C until used.

PG12

The methods used for PG12 were essentially the same as for PG1 with 10 the following exceptions. The 5' oligonucleotide primer sequence was GCGCGTATACATGAATAGCAGACATCTGACAATCACAATCATTGCCGG, the 3' oligonucleotide primer sequence was GCGCAGATCTGCTGTTCTGTGAGTGCAGTTGTTTAAGTG and DNA was amplified with Tli DNA polymerase. The PCR product was cloned into pCR-15 Blunt and transformed into E. coli Top10F'cells before subcloning into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into E. coli BL21. Purification of the recombinant protein was essentially the same as PG11 except 0.5% DHPC 20 (1,2-Diheptanoyl-sn-glycero-3-phosphocholine; Avanti) in 50mM Tris, 50mM NaCl, pH8.0 was used to solubilise the inclusion bodies instead of sarkosyl and the DHPC was diluted to 0.1% before addition to the Ni-NTA and 0.1% DHPC was added to all buffers.

25 **PG13**

The methods used for PG13 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCCATATGCGGACAAAAACTATCTTTTTTGCG, the 3' oligonucleotide primer sequence was

30 GCGCCTCGAGGTTGTTGAATCGAATCGCTATTTGAGC and DNA was amplified with Tli DNA polymerase. The PCR product was cloned the expression plasmid pET24b using the Nde I and Xho I restriction sites and transformed into *E. coli* BL21. Purification of the recombinant protein was essentially the same as PG3 using 6M urea and 1% NOG (n-octyl glucoside;

35 Sigma) was added to the dialysis buffer. Removal of urea was not proceeded

past 2M urea as the protein was insoluble at lower concentrations of urea. Purified protein was stored at 4°C until required.

PG14

5

10

20

25

30

The methods used for PG12 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGCGCCATGACGGACAACAACAACGTAATATCG, the 3' oligonucleotide primer sequence was GCGCCTCGAGTTACTTGCGTATGATCACGGACATACCC and DNA was amplified with Tli DNA polymerase. The PCR product was cloned the expression plasmid pProEx-1 using the Ehe I and Xho I restriction sites and transformed into *E. coli* BL21. Purification of the recombinant protein was essentially the same as PG12.

15 PG15

The methods used for PG15 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CAAAAGTATACTAATAAATATCATTCTCAA, the 3' oligonucleotide primer sequence was GCTTATGGTACCTTTGGTCTTATCTATTAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Kpn I restriction sites and transformed into *E. coli* ER1793.

PG22

The methods used for PG22 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was CCCCGGATCCGATGCGACTGATCAAGGC, the 3' oligonucleotide primer sequence was CCCCCTCGAGCGGAACGGGGTCATAGCC and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was cloned into the expression plasmid pET24b using the Bam HI and Xho I restriction sites and transformed into *E. coli* BL21DE3. Once PG22 was purified dialysis was performed in the same manner as for PG1 but in the presence of 1M imidazole.

50

PG24

5

10

15

20

25

The methods used for PG24 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACATGAATTACCTGTACATAC, the 3' oligonucleotide primer sequence was CGCGGGATCCGTTCGATTGGTCGTCGATGG and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was digested with Bst Z171 and Bam HI and ligated into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* ER1793. Due to the low level of expression of PG24 purification was not proceeded with except on small scale.

PG24A

A modified version of PG24 was also cloned and expressed. PG24A is the same as PG24 with the predicted N-terminal sequence removed. The methods used for PG24A were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGCATATGGAGATTGCTTTCCTTTCTTCG, the 3' oligonucleotide primer sequence was CGCGCTCGAGTTAGTTCGATTGGTCGTCG and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was cloned into the expression plasmid pProEx-1 using the Nde I and Xho I restriction sites and transformed into *E. coli* ER1793. Purification of the recombinant protein was essentially the same as PG3 except 8M urea was used to solubilise the inclusion bodies and in the buffers used for the Ni-NTA column purification. Urea was removed by sequential dialysis (4M then 2M, then 1M then 0.5M then 0M urea all in 50mM Tris, 500mM NaCl, 8% glycerol, pH7.4). Purified protein was stored frozen at -80°C until required.

PG29

The methods used for PG29 were essentially the same as for PG3 with
the following exceptions. The 5' oligonucleotide primer sequence was
GCGCGATATCGCTAGCATGAAAAAGCTATTTCTC, the 3' oligonucleotide
primer sequence was
GCGCAGATCTCTCGAGTTTGCCATCGGATTGCGGATTG and DNA was
amplified with Pfu DNA polymerase being used. The PCR product was
cloned into pCR-Blunt (InVitrogen) and transformed into E. coli
Top10F'before subcloning into the expression plasmid pGex-stop RBS(IV)

using the EcoR V and Bgl II restriction sites and transformed into *E. coli* BL21. 6M urea was used throughout the purification process.

PG30

5

10

15

20

25

The methods used for PG30 were essentially the same as for PG3 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TACGGAATTCGTGACCCCCGTCAGAAATGTGCGC, the 3' oligonucleotide primer sequence was

CTATGCGGCCGCTTTGATCCTCAAGGCTTTGCCCGG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into $E.\ coli\ BL21DE3$. Expression studies and immunoreactivity studies were carried out on whole $E.\ coli\ lysates$ of PG30. 10ml cultures of recombinant $E.\ coli\ were$ grown to an OD of 2.0 (A_{600nm}) in terrific broth and the cells were induced with 0.5mM IPTG and samples taken for analysis at 4 hours post induction. Purification was not done for these studies.

PG31

The methods used for PG31 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was CGGGGAATTCGCAAAAATCAATTTCTATGCTGAA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTGTATGCAATAGGGAAAGCTCCGA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E.coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

30 PG32

The methods used for PG32 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGCAGAATTCCAGGAGAATACTGTACCGGCAACG, the 3' oligonucleotide primer sequence was CTATGCGGCCGCCTTGGAGCGAACGATTACAACAC and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG33

5

10

15

25

30

35

The methods used for PG33 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein.

The 5' oligonucleotide primer sequence was TGCAGAATTCCAAGAAGCTACTACACAGAACAAA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTTCCGCTGCAGTCATTACTACAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

20 PG35

The methods used for PG35 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGAATTCATGAAACAACTAAACATTATCAGC, the 3' oligonucleotide primer sequence was GCGTGCGGCCGCGAAATTGATCTTTGTACCGACGA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG36

The methods used for PG36 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was AAAGGAATTCTACAAAAAGATTATTGCCGTAGCA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCGAACTCCTGTCCGAGCACAAAGT and DNA was amplified with the Tth XL PCR kit. The PCR product was

cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunore activity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

5

10

15

20

25

PG37

The methods used for PG37 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was TGGCGAATTCAAACGGTTTTTGATTTTGATCGGC, the 3' oligonucleotide primer sequence was CTATGCGGCCGCCTTGCTAAAGCCCATCTTGCTCAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG38

The methods used for PG38 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CCTCGAATTCCAAAAGGTGGCAGTGGTAAACACT, the 3' oligonucleotide primer sequence was CTATGCGGCCGCCTTGATTCCGAGTTTCGCTTTTAC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

30 PG39

The methods used for PG39 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCTGGATCCCAAGGCGTCAGGGTATCGGGCTAT, the 3' oligonucleotide primer sequence was CTATGCGGCCGGAATTCGACGAGGAGACGCAGGT and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG40

5

10

15

25

30

35

The methods used for PG40 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCTGAATTCAAGACGGACAACGTCCCGACAGAT, the 3' oligonucleotide primer sequence was CTATGCGGCCGCAAGTTGACCATAACCTTACCCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

20 **PG41**

The methods used for PG41 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GACTGAATTCCAAAACGCCTCCGAAACGACGGTA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTTGTTCGGGAATCCCCATGCCGTT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG42

The methods used for PG42 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was GTTTGAATTCGCAAATAATACTCTTTTGGCGAAG, the 3' oligonucleotide

primer sequence was

GAGTGCGGCCGCTTTGCCGGACATCGAAGAGATCGTC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG43

5

10

15

The methods used for PG43 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGAATTCAAAAAAGAAAACTTTGGATTGCG, the 3' oligonucleotide primer sequence was CTATGCGGCCGCCTTCAAAGCGAAAGAAGCCTTAAC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

20 PG44

The methods used for PG44 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCCGAATTCTGTAAGAAAAATGCTGACACTACC, the 3' oligonucleotide primer sequence was CTATGCGGCCGCCTTTTTCCCGGGCTTGATCCCGAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG45

The methods used for PG45 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer

sequence was GACAGGATCCTGCTCCACCACAAAGAATCTGCCG, the 3' oligonucleotide primer sequence was CTATGCGGCCGCGAAGGGATAGCCGACAGCCAAAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10 PG46

The methods used for PG46 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CTCGGAATTCCGTTATGTGCCGGACGGTAGCAGA, the 3' oligonucleotide primer sequence was CTATGCGGCCGGAACGGATAGCCTACTGCAATGT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG47

The methods used for PG47 were essentially the same as for PG30
with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was
CGCCGAATTCCAAACAGTGGTGACCGGTAAGGTGATCGATTCAGAA, the
3' oligonucleotide primer sequence was
CTATGCGGCCGCGAAGTTTACACGAATACCGGTAGACCAAGTGCGGCC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies

and immunoreactivity studies were carried out on whole E. coli lysates.

35 Purification was not done for these studies.

PG48

5

10

15

20

25

30

35

The methods used for PG48 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAAAAATCCAAGCAGGTACAGCGA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTCGTAACCATAGTCTTGGGTTTTG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG49

The methods used for PG49 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GAACGGATCCAACGAGCCGGTGGAAGACAGATCC, the 3' oligonucleotide primer sequence was

GAGTGCGGCCGCTAATCTCGACTTCATACTTGTACCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG50

The methods used for PG50 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GCTGGGATCCGCGACAGACACTGAGTTCAAGTAC, the 3' oligonucleotide primer sequence was CTATGCGGCCGCAACTTCACTACCAAGCCCATGT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG51

5

10

15

20

25

30

35

The methods used for PG51 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TCTTGAATTCGCGCAAAGTCTTTTCAGCACCGAA, the 3' oligonucleotide primer sequence was

CTATGCGGCCGCACTTTTTCGTGGGATCACTCTCTT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG52

The methods used for PG52 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was AGAAGAATTCAAACGGACAATCCTCCTGACGGCA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCGAAGTCTTTGCCCTGATAGAAATC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG53

The methods used for PG53 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCTGAATTCGCGAATCCCCTTACGGGCCAATCG, the 3' oligonucleotide primer sequence was CTATGCGGCCGGTCCGAAAGGCAGCCGTAATAGG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and

transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

5 **PG54**

10

15

The methods used for PG54 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGCTGAATTCCAGATTTCGTTCGGAGGGGAACCC, the 3' oligonucleotide primer sequence was CTATGCGGCCGCCTGCTTCACGATCTTTTGGCTCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG55

with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGAGGGATCCGAGCTCTCTATTTGCGATGGCGAG, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTCTTACCTGACTTCTTGTCACGAAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG56

30

35

The methods used for PG56 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was AAATGGATCCCGAAAAATTTTGAGCTTTTTGATG, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTTTGATTCGTAATTTTTCCGTATC and DNA was amplified with the Tth XL PCR kit. The PCR product was

cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

5

10

15

20

25

30

35

PG57

The methods used for PG57 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGGATCCCAAGAGATCTCAGGCATGAATGCA, the 3' oligonucleotide primer sequence was GAGTGCGGCCTCGGCCTCTTTATCTCTACCTTTTC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG58

The methods used for PG58 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGGTGAATTCCAAACCCCACGAAATACAGAAACC, the 3' oligonucleotide primer sequence was

GAGTGCGGCCGCTGAAAGTCCAGCTAAAACCGGCGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG59

The methods used for PG59 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAACAAGAGAAGCAGGTGTTTCAT, the 3'

oligonucleotide primer sequence was GAGTGCGGCCGCTGAAGATGCTCTTATCGTCCAAACG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG60

5

15

20

The methods used for PG60 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGGAATTCCAGATGCTCAATACTCCTTTCGAG, the 3' oligonucleotide primer sequence was

GAGTGCGCCGCTGAAGAGGTAGGAGATATTGCAGAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG61

The methods used for PG61 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCAGAATTCCCCGTCTCCAACAGCGAGATAGAT, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGAAATCGATTGTCAGACTACCCAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

62

PG62

The methods used for PG62 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGCGGTTTCCGATGGTGCAGGGA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGAAGTGAAATCCGACACGCAGCTG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG63

10

15

20

25

30

35

The methods used for PG63 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAAGAAGCAAACACTGCATCTGAC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGAAAGTGTACGCAACACCCACGCC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and

transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG64

The methods used for PG64 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGAGTCGTCCTGCTCTTAGACTG, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGAAGCGAACACCGAGACCCACAAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG65

5

10

15

20

25

The methods used for PG65 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGGATCCATCGGACAAAGCCGCCCGGCACTT, the 3' oligonucleotide primer sequence was

GAGTGCGCCGCTAAAGCGGTAACCTATGCCCACGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG66

The methods used for PG66 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GTTTGAATTCCAAGACGTTATCAGACCATGGTCA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTAAAATGAGTGGAGAGCGTGGCCAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

30 **PG67**

The methods used for PG67 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GAACGAGCTCGCGGAACGTCCTATGGCCGGAGCA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTATACCAAGTATTCGTGATGGGACG and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Sac I and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG68

5

10

15

25

30

35

The methods used for PG68 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GCTTGCGGCCGCCCTTATGAAAGATTTGCAGAT, the 3' oligonucleotide primer sequence was GGTGCTCGAGTATACTCAACAAGCACCTTATGCAC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Not I and Xho I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

20 PG69

The methods used for PG69 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGGAAGGGGAGGGGAGTGCCCGA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGAAGCTGTAGCGGGCTTTGAACCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG70

The methods used for PG70 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer

sequence was CGGTGGATCCTCGCAAATGCTCTTCTCAGAGAAT, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTAAACGAAATATCGATACCAACATC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10 PG71

5

15

20

The methods used for PG71 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGAACAATACCCTCGATGTACAC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTATTGCCGGTAGGATTTCCTTGTCC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG72

with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCGGAGAGCGACTGGAGACGGACAGC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTATGATTGCCTTTCAGAAAAGCTAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed TGCTGAATTCGGAGAGCGACTGGAGACGGACAGC into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG73

The methods used for PG73 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGGTGAATTCCAACAGACAGGACCGGCCGAACGC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTTAAGAAAGGTATCTGATAGATCAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG74

10

15 The methods used for PG74 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAAGAAAATAATACAGAAAAGTCA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGAGGTTTAATCCTATGCCAATACT and DNA was 20 amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG75

25

30

35

The methods used for PG75 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGGGATCCGCTCAGGAGCAACTGAATGTGGTA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGTGGAACAAATTGCGCAATCCATC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG76

5

10

15

20

25

The methods used for PG76 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCAGAATTCGGAAACGCACAGAGCTTTTGGGAA, the 3' oligonucleotide primer sequence was

GAGTGCGCCGCTTACCTGCACCTTATGACTGAATAC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG77

The methods used for PG77 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAAGAGAAAAAGGATAGTCTCTCT, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTCTTCTTATCGCCATAGAATACAGG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

30 PG78

The methods used for PG78 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAGGATTCTTCCCACGGTAGCAAT, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTATCATGATAGTAAAGACTGGTTCT and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG79

5

10

15

The methods used for PG79 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was TGCTGAATTCGTAGTGACGCTGCTCGTAATTGTC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGCCTGCCTTTCTGCCTGACG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG80

with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAATTCCAAAACGTGCAGTTGCACTACGAT, the 3' oligonucleotide primer sequence was

GAGTGCGCCCGCTGTTGAAAGTCCATTTGACCGCAAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done

PG81

for these studies.

30

35

The methods used for PG81 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GTTTGAATTCCAGGATTTTCTCTATGAAATAGGA, the 3'

oligonucleotide primer sequence was GAGTGCGGCCGCTTTGTTTATTACAAAAAGTCTTACG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG82

5

15

20

25

30

The methods used for PG82 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GAACGAATTCCAGAACAACAACTTTACCGAGTCG, the 3' oligonucleotide primer sequence was

GAGTGCGCCGCTGTTCAGTTTCAGCTTTTTAAACCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG84

The methods used for PG84 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGGATCCCAGAATGATGACATCTTCGAAGAT, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTATTGCGTCCCCGGCCACTACGTCC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG85

5

10

15

20

25

30

35

The methods used for PG85 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGGTGAATTCGTACCAACGGACAGCACGGAATCG, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTCAGATTGGTGCTATAAGAAAGGTA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG86

The methods used for PG86 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGGATCCCAAACGCATGATCATCTCATCGAA, the 3' oligonucleotide primer sequence was

GAGTGCGGCCGCTGTGGTTCAGGCCGTGGGCAAATCT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG87

The methods used for PG87 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGGAATTCCAGAGCTATGTGGACTACGTCGAT, the 3' oligonucleotide primer sequence was GAGTGCGCCGCTATTACTGTGATTAGCGCGACGCTG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG88

5

10

15

20

25

The methods used for PG88 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCAGAATTCGCCGAATCGAAGTCTGTCTCTTTC, the 3' oligonucleotide primer sequence was

GAGTGCGGCCGCTCGGCAAGTAACGCTTTAGTGGGGA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG89

The methods used for PG89 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAATCGAAGTTAAAGATCAAGAGC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTATTAGTCCAAAGACCCACGGTAAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

30 **PG90**

The methods used for PG90 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAAACAACGACGAACAGTAGCCGG, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTTTTTTGTTGTGATACTGTTTGGGC and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG91

10

15

25

30

35

The methods used for PG91 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGACGATGGGAGGAGGAGATGATGTC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTTTCCACGATGAGCTTCTCTACGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

20 **PG92**

The methods used for PG92 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAATTCGCCGATGCACAAAGCTCTGTCTCT, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTTCGAGGACGATTGCTTAGTTCGTA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG93

The methods used for PG93 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer

sequence was GGCCGAGCTCCAAGAGGAAGGTATTTGGAATACC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGCGAATCACTGCGAAGCGAATTAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Sac I and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10 PG94

5

15

20

The methods used for PG94 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAGCTCCAAGAGGAAGGTATTTGGAATACC, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTTTGTCCTACCACGATCATTTTCTT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG95

PG96

5

10

15

20

25

30

35

The methods used for PG96 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAGCTCCAAACGCAAATGCAAGCAGACCGA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTTTTGAGAATTTTCATTGTCTCACG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Sac I and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG97

The methods used for PG97 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGGGATCCCAGTTTGTTCCGGCTCCCACCACA, the 3' oligonucleotide primer sequence was

GAGTGCGGCCGCTCTGTTTGATGAGCTTAGTGGTATA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG98

The methods used for PG98 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCAGAATTCCAAGAAAGAGTCGATGAAAAAGTA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTTAGCTGTGTAACATTAAGTTTTTATTGAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and

immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG99

5

The methods used for PG99 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCAAGGACAATTCTTCTTACAAACCT, the 3' oligonucleotide primer sequence was

GAGTGCGGCCGCTTCGAATCACGACTTTTCTCACAAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG100

20

25

The methods used for PG100 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAGTCTTTGAGCACAATCAAAGTA, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTGATAGCCAGCTTGATGCTCTTAGC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

30 **PG101**

The methods used for PG101 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was TGCTGAATTCAAAGGCAAGGGCGATCTGGTCGGG, the 3' oligonucleotide primer sequence was

35 GAGTGCGGCCGCTTCTCTCTCGAACTTGGCCGAGTA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the

expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

5

10

15

PG102

The methods used for PG102 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAATTCCAGATGGATATTGGTGGAGACGAT, the 3' oligonucleotide primer sequence was GAGTGCGGCCGCTCTCTACAATGATTTTTTCCACGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

PG104

20

The methods used for PG104 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GAACGGATCCAACGTGTCTGCTCAGTCACCCGA, the 3' oligonucleotide primer sequence was

GAGTGCGGCCGCTTCTGAGCGATACTTTTGCACGTAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into E. coli BL21DE3. Expression studies and immunoreactivity studies were carried out on whole E. coli lysates. Purification was not done for these studies.

Animal antisera and human patient sera.

Various antisera were raised for detecting the expression and refolding of the recombinant *P. gingivalis* proteins. A whole cell antisera was raised by injecting New Zealand White rabbits with 3 doses of sonicated

15

20

25

30

35

P. gingivalis (strain W50) containing approximately 2mg of protein. The first dose was given in Freunds complete adjuvant (FCA) and the second and third doses were given in Freunds incomplete adjuvant (IFA) at 3 week intervals. Doses (1ml) were given intramuscularly into the hind legs and rabbits bled 7 days after the last dose, the blood clotted and serum removed and stored at -20°C until required. A second rabbit antisera was produced in a similar manner but using a sarkosyl insoluble fraction (each dose was 0.69mg of protein) derived from P. gingivalis W50 according to the method of Doidg and Trust T. et al 1994 as the immunogen. A third rabbit antisera was produced in a similar manner to the first only the sarkosyl soluble fraction (1mg of protein per dose) derived from P. gingivalis W50 cells according to the method of Doidg P. and Trust TJ. (1994 Infect Immun 62:4526-33) was used as the immunogen.

A "protected rat serum" pool was also used in these studies and was obtained from rats immunised with formalin killed whole *P. gingivalis* cells in FIA (strain ATCC 33277; 2 doses of $2x10^9$ cells, 3 weeks apart). Rats were then challenged 2 weeks after their last dose with live *P. gingivalis* cells (strain 33277) given orally as previously described (Klaussen B. et al. 1991, Oral Microbiol Immunol 6:193-201) and the serum obtained from these rats 6 weeks after the final challenge inoculation at the time of sacrifice.

Human sera were obtained from adult patients undergoing treatment or assessment for periodontitis at an outpatient clinic. These patients had at least 6 teeth with 6mm attachment loss and had *P. gingivalis* present in their sub-gingival plaque as detected using a *P. gingivalis* specific DNA probe. Sera was pooled from these patients and compared to a pool of sera from periodontally healthy patients.

Immunization and Murine Lesion Model Protocols

The mouse abscess model was used to assess the efficacy of immunising mice with recombinant *P. gingivalis* proteins in protecting mice from formation of a subcutaneous abscess. This model has been used by others as a predictor of potential vaccines against periodontal disease (Bird PS, et al. 1995 J. Periodontol. 66:351-362. BALB/c mice 6-8 weeks old were immunised by subcutaneously injecting them with 0.1 ml containing either 10 or 20µg of recombinant *P. gingivalis* protein, 20µg of *E. coli* lysate protein,

2 x 10° formalin killed cells of *P. gingivalis* strain 33277 emulsified in incomplete Freund's adjuvant (IFA; Sigma) on day 0. At day 21 mice were re-injected with the same dose and then bled 1 week later and evaluated for antibody levels. At day 35 mice all mice were challenged with approximately 2 x 10° cells of live *P. gingivalis* (ATCC 33277) by subcutaneous injection in the abdomen. Following challenge mice were monitored daily for weight loss and the size of the lesion measured for the next 10 days. Lesion sizes were measured by length and width and expressed as mm². Groups were statistically analysed using a Kruskal-Wallis one-way ANOVA and were also individually examined using the unpaired t test or Mann-Whitney rank sum test using the Instat statistical package.

Figure 1 shows the results of one experiment at day 4 after challenge (lesions were at maximum size at this time point). Control mice immunised with $E.\ coli$ lysate showed large lesions while mice immunised with killed cells of $P.\ gingivalis$ strain 33277 were fully protected. This indicates that whole cells provide protection against $P.\ gingivalis$ while $E.\ coli$ protein immunised mice were not protected. Mice given the various PG recombinant proteins showed significant levels of protection for PG2, PG22, PG24 and PG29 (p<0.05 unpaired t test) while PG8A was not quite significantly different (p=0.07) compared to the $E.\ coli$ control group.

Figure 2 shows the results of a separate experiment using combinations of recombinant proteins. Mice given PG1 + PG2 showed a significant level of protection compared to control mice give E. coli lysate (p<0.026 unpaired t test).

Immunoscreening

Cloned candidates were cultured in 15ml of Terrific broth, induced with IPTG and sampled at 4h post-induction. One ml of culture was removed, pelleted and the cells resuspended in a volume of PBS determined by dividing the OD $A_{600\mathrm{nm}}$ of the culture by 8. An aliquot of lysate (100µl) was added to 100µl of 2x sample reducing buffer (125mM Tris pH 6.8, 20% glycerol, 4% SDS, 80mM DTT, 0.03% bromophenol blue) and boiled for 10min. SDS-PAGE was performed according to the method of Laemmli UK. 1970 (Nature 227:680-685) using 4-20% 1.0mm Tris-Glycine gels (Novex) according to the manufacturers recommendations. Proteins were transferred

5

10

15

20

25

30

10

15

20

25

onto Hybond-C Extra nitrocellulose membranes (Amersham) by transblotting and the membranes were then blocked for 2h at room temperature (RT) in 5% skim milk in 20mM Tris, 0.5M NaCl, 0.05% Tween-20, pH 7.5 (TTBS).

Immunoscreening was performed separately with the rabbit anti-P. gingivalis whole cell serum, the rat protective serum, a pool of human periodontal patients serum, and in many cases an anti-T7-Tag antibody HRP conjugate (Novagen). Prior to use, the rabbit, rat and human sera were diluted 1/5000, 1/1000 and 1/500 respectively in 5% skim milk in TTBS and absorbed with 100µl (for the rabbit serum) or 250µl (for the rat and human sera) E. coli extract (20mg/ml; Promega) for 6h at RT.

Membranes were incubated overnight at RT with the absorbed antisera, or for 1 hr at RT with 1/5000 diluted anti-T7-Teg conjugate. Following 3x10min washes with TTBS, HRP-conjugated anti-rabbit (Silenus), anti-mouse (Silenus) or anti-human (KPL) antibody, diluted 1/5000 in 5% skim milk in TTBS, was added for 1h at RT. Membranes were washed as before, prior to addition of TMB membrane peroxidase substrate (KPL) for detection of immunoreactive proteins. Results of reactivity for the recombinant *P. gingivalis* proteins is shown in Table 7.

In addition some of the sera (pooled sera diluted 1/1000) from the mice immunised with *P. gingivalis* recombinant proteins (prior to challenge) were analysed for their reactivity against Western blots of whole native W50 *P. gingivalis* proteins using similar techniques as those outlined above. PG2, PG8A, PG29 and PG3 all showed bands at a similar molecular weight to that of the recombinant PG protein in the native W50 blot. This indicates that PG proteins are expressed in the W50 strain and that the recombinant proteins have at least some identical immunogenicity to the native proteins.

m-RNA analysis

30 Hot Phenol RNA Extraction

P. gingivalis W50 cells (150ml culture) were grown anaerobically to mid log phase (OD A_{600} =0.18) mixed with 50% glycerol and stored at -70°C until RNA extraction. Cells were pelleted by centrifugation at 6000g, and resuspended in 8ml ASE (20mM NaOAc, 0.5% SDS, 1mM EDTA). An equal volume of 20mM NaOAc(pH 4.5)-saturated phenol was added and mixed by

WO 99/29870 PCT/AU98/01023

80

shaking for 30 seconds, incubated at 65°C for 5 minutes, followed by a further 5 second shaking and repeated incubation. After cooling, 2ml chloroform was added and mixed by shaking for 5 seconds, and the mixture spun at 10000g for 10 minutes at 4°C. The top aqueous phase was transferred and re-extracted by repeating the phenol and chloroform steps. The aqueous phase was transferred again and 100U RNase inhibitor (RNAsin; Promega) were added. RNA was precipitated with 3 volumes 100% ethanol at -20°C overnight. The RNA precipitate was recovered by centrifugation at 10000g at 4°C for 15 minutes, then washed with 100% ethanol, dried and resuspended in 600μ l sterile, deionised, dH₂O with 1μ l of fresh RNase inhibitor. RNA was aliquoted and stored at -70°C. The RNA concentration was determined spectrophotometrically. A formaldehyde RNA gel confirmed RNA integrity (Sambrook J. et al. 1989, Molecular Cloning. A laboratory manual. Cold Spring Laboratory Press, New York. 2nd Edition).

15

20

25

30

35

10

5

RT-PCR

The isolated RNA was used as a template for Reverse Transcription (RT) to produce cDNA. Varying RNA concentrations were used for the RT as each RNA transcript was potentially present at different levels. Subsequent amplification of the cDNA was performed using Polymerase Chain Reaction (PCR). RT-PCR was performed using GeneAmp® RNA PCR Kit (Perkin Elmer) according to the manufacturer's protocol with the following exception to the PCR; 35 cycles were performed as follows: Melt phase 95°C for 30 seconds, Anneal phase varied between 50-60°C for 30 seconds, Extension phase 72°C for 1 minute. Amplification was performed in a PTC-100 Programable Thermal Controller (MJ Research Inc.). As a control to demonstrate that the amplified product did not arise from contaminating DNA, Reverse Transcriptase (RTase) was omitted from a parallel tube. The PCR products were examined against DNA markers (GIBCO 1kB ladder) on a 1% agarose gel stained with ethidium bromide.

RT-PCR results are shown in Table 6 using the oligonucleotide primers as used in "Cloning, expression and purification of recombinant *P. gingivalis* genes" section described above, except for the following changes. For PG1 the 3' reverse primer used was

GCGCCTCGAGATTCATTTCCTTATAGAG, for PG4 the 5' forward primer was CTTCTTGTCGACTACAGCGGACATCATAAAATC and the 3' reverse primer was TTCCACCTCGAGTTAACGCAACTCTTCTTCGAT, for PG6 the 5' forward primer was TAAAGAATTCTGCCTCGAACCCATAATTGCTCCG, for PG10 the 5' forward primer was CGCGCATATGGATAAAGTGAGCTATGC and the 3' reverse primer was CGCGCTCGAGTTTGTTGATACTCAATAATTC, for PG13 the 5' forward primer was GCCCGGCGCCCATGCGGACAAAAACTATCTTTTTTGCG and the 3' reverse primer was

of *P. gingivalis* transcripts is a likely indication that RNA for a specific candidate is present and that the protein is produced. However, where there is no amplification achieved this does not indicate that this gene is never transcribed and may be the result of the culture conditions or the state of the cells when harvested.

Table 6. Expression of PG m-RNA with *in vitro* grown *P. gingivalis* W50. The symbols are + band visible on agarose gel, - no band present on agarose gel, ND not detected.

| PG # | RNA | Annealing | RT-PCR | PCR (-RT) | Approx. | Expected |
|------|------|-----------|--------|-----------|----------|----------|
| | μg | temp. °C | | | fragment | fragment |
| | | | | | size bp | size bp |
| 1 | 0.15 | 55 | + | - | 1300 | 1362 |
| 2 | 1.0 | 50 | + | - | 3200 | 3051 |
| 3 | 0.15 | 60 | + | - | 720 | 690 |
| 4 | 2.9 | 55 | - | - | N.D. | 2000 |
| 5 | 0.02 | 50 | + | - | 1000 | 947 |
| 6 | 1.0 | 55 | + | | 1000 | 972 |
| 8A | 0.15 | 50 | + | <u>-</u> | 1200 | 1278 |
| 10 | 0.15 | 55 | + | - | 590 | 585 |
| 11 | 0.10 | 60 | + | - | 960 | 942 |
| 12 | 0.02 | 60 | + | - | 880 | 831 |
| 13 | 1.0 | 50 | + | - | 2150 | 2274 |
| 14 | 0.15 | 60 | + | - | 1050 | 996 |

| PG# | RNA µg | Annealing temp. °C | RT-PCR | PCR (-RT) | Approx. fragment size bp | Expected fragment size bp |
|-----|-----------|--------------------|--------|-----------|--------------------------|---------------------------|
| 22 | 1.0 | 60 | - | - | N.D. | 228 |
| 24 | 1.0 | 55 | + | + | 1150 | 1194 |
| 29 | 0.15 | 60 | + | - | 880 | 885 |

Table 7: Immunoblot results of proteins expressed in *E.coli* against rabbit, rat and human antisera. Deduced MW was calculated from amino acid sequence of the *P. gingivalis* proteins, some of which had their N-terminal signal sequences removed. Apparent MW was determined from SDS-PAGE gels. The N- and C-terminal tags add approximately 2.5 KDa to the deduced MW of the recombinant proteins. The symbols are + positive, - negative, +/- weak positive, ND not done.

| Protein number | Deduced MW (KDa) | Apparent MW (KDa) | Antisera reactivity | | | |
|-------------------|---------------------|----------------------|---------------------|--------|-----|-------|
| | | | T 7 | Rabbit | Rat | Human |
| PG1 | 47.5 | 63 | ND | - | - | - |
| PG2 | 112.4 | 125.7 | ND | + | - | - |
| PG3 | 22.6 | 18.3 | ND | -a | - | - |
| PG4 | 7 5 | 90.6 | ND | - | - | - |
| PG5 | 34.9 | 43.8 | ND | - | - | - |
| PG6 | 36.7 | 47.1 | ND | - | - | - |
| PG8 | 67.5 | 63.1 | ND | _b | - | - |
| PG8A | 47.7 | 90.6 | ND | - | - | - |
| PG10 | 21.3 | 25.5 | ND | + | - | + |
| PG11 | 36.2 | 42.4 | ND | - | - | _ |
| PG12 | 30. <i>7</i> | 30.6 | ND | - | - | - |
| PG13 | 84.5 | 101 | ND | - | _ | - |
| PG14 | 36 | 42.4 | ND | - | + | + |
| PG22 | 8.6 | 11.1 | . ND | - | - | _ |
| PG24A | 47 | 63.1 | ND | - | • | - |
| PG29 | 31.1 | 40.9 | ND | + | + | + |

| Protein number | Deduced MW (KDa) | Apparent MW (KDa) | Antisera reactivity | | | | |
|-------------------|---------------------|----------------------|---------------------|--------|-----|-------|--|
| number | WW (KDa) | MW (KDa) | T 7 | Rabbit | Rat | Human | |
| PG30 | 35.1 | 46.9 | + | - | _ | - | |
| PG31 | 16.7 | - | _ | - | ** | - | |
| PG32 | 41.2 | 59.5 | + | + | + | - | |
| PG33 | 39.9 | 52.7 | + | + | + | - | |
| PG35 | 92.6 | 116.6 | + | - | - | - | |
| PG36 | 98.9 | 120.2 | _ | _ | - | - | |
| PG37 | 18.8 | 23.1 | + | + | - | - | |
| PG38 | 16.1 | 22.9 | + | - | _ | - | |
| PG39 | 87.9 | 116 .6 | + | - | _ | - | |
| PG40 | 76.6 | 103.1 | + | - | _ | - | |
| PG41 | 48.3 | 81.1 | + | - | + | + | |
| PG42 | 59.3 | 73.9 | + | - | - | - | |
| PG43 | 27.1 | 50.3 | + | - | - | - | |
| PG44 | 28.6 | 32.3 | + | - | + | - | |
| PG45 | 84 | 100.6 | + | - | - | - | |
| PG46 | 83 | 97.7 | + | - | - | - | |
| PG47 | 93.7 | 42.5 | + | + | - | + | |
| PG48 | 45.2 | 37.9 | + | - | - | _ | |
| PG49 | 33.3 | 64.1 | + | - | + | - | |
| PG50 | 91.9 | 113.2 | + | + | _ | - | |
| PG51 | 19.6 | 27.2 | + | - | _ | _ | |
| PG52 | 50.4 | 64.4 | + | + | _ | + | |
| PG53 | 47.4 | 45.4 | + | - | _ | + | |
| PG54 | 101.4 | 46.7 | + | + | _ | - | |
| PG55 | 70.4 | 68.4 | + | - | _ | | |
| PG56 | 142.3 | - | - | - | - | - | |
| PG57 | 100 | 134.5 | + | + | + | + | |
| PG58 | 63 | 82.9 | + | - | _ | - | |
| PG59 | 33.3 | 43.6 | + | _ | - | - | |
| PG60 | 55.6 | 77.8 | + | _ | _ | - | |
| PG61 | 81.5 | 107.3 | + | - | - | - | |

| Protein number | Deduced MW (KDa) | Apparent MW (KDa) | | Antisera reactivity | | | | |
|-------------------|---------------------|----------------------|-----|---------------------|--------------------------------|-------|--|--|
| | | | Т7 | Rabbit | Rat | Human | | |
| PG62 | 51.9 | 58.4 | + | - | - | - | | |
| PG63 | 29.6 | 43.6 | + | - | - | - | | |
| PG64 | 18.5 | 26.9 | + | - | - | - | | |
| PG65 | 25.9 | 28.8 | + | - | - | - | | |
| PG66 | 22.2 | 25.1 | + | + | - | - | | |
| PG67 | 103.7 | 105 | + | - | - | - | | |
| PG68 | 133.3 | 30.7 | + | - | + | + | | |
| PG€9 | 44.4 | 50.8 | + | - | - | - | | |
| PG70 | 25.9 | 30.8 | + | - | - | - | | |
| PG71 | 88.9 | 105 .5 | + | - | - | - | | |
| PG72 | 40.7 | 49.8 | + | - | - | - | | |
| PG73 | 40.7 | 29 | +/- | - | - | - | | |
| PG74 | 22.2 | 32.5 | + | _ | - | - | | |
| PG75 | 40.7 | 46.7 | + | - | _ | - | | |
| PG76 | 48.1 | 55.6 | + | - | - | + | | |
| PG77 | 29.6 | 36.9 | + | - | _ | _ | | |
| PG78 | 33.3 | 35.4 | + | - | - | - | | |
| PG79 | 33.3 | - | _ | - | - | - | | |
| PG80 | 25.9 | 20.5 | + | - | _ | - | | |
| PG81 | 23 | 25.8 | + | - | - | - | | |
| PG82 | 44.8 | 48.5 | + | - | - | - | | |
| PG84 | 41.7 | 52.4 | + | - | _ | +/- | | |
| PG85 | 62.7 | 72.4 | + | - | _ | - | | |
| PG86 | 21.7 | 27.4 | + | - | _ | +/- | | |
| PG87 | 83 | 91.3 | + | - | ****************************** | + | | |
| PG88 | 27 | 40.1 | + | - | - | _ | | |
| PG89 | 26.2 | 29.4 | + | - | - | - | | |
| PG90 | 23 | 28.4 | + | - | - | - | | |
| PG91 | 57.2 | 85. <i>7</i> | . + | + | + | + | | |
| PG92 | 83.6 | 110.4 | + | - | - | + | | |
| PG93 | 83.4 | 110.4 | + | - | - | + | | |

| Protein number | Deduced MW (KDa) | Apparent MW (KDa) | Antisera reactivity | | | |
|-------------------|---------------------|----------------------|---------------------|--------|-----|-------|
| | | | T7 | Rabbit | Rat | Human |
| PG96 | 59.3 | 70.3 | + | + | + | + |
| PG97 | 44.4 | 5 <i>7</i> .5 | + | - | + | + |
| PG98 | 33.3 | 36 | + | - | _ | - |
| PG99 | 40.7 | 55.6 | + | - | + | + |
| PG100 | 29.6 | 10.8 | + | - | - | - |
| PG101 | 14.8 | 19.7,14.1 | + | - | - | - |
| PG102 | 59.3 | 70.3 | + | - | _ | + |
| PG104 | 40.7 | 57.5 | + | - | - | + |

- a. Positive reaction detected with the rabbit antiserum to sarkosyl insoluble *P. gingivalis* antigen.
- **b.** Purified protein demonstrated weak positive reaction with the rabbit antiserum to whole *P. gingivalis*.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

WO 99/29870 PCT/AU98/01023

References.

10

1. Lipman DJ, Pearson WR. 1985. Rapid and sensitive protein similarity searches. Science 277:1435-1441.

- 5 2. Horton, P. and Nakai, K. (1996). A probabilistic classification system for predicting the cellular localization sites of proteins. Intellig. Syst. Mol. Biol. 4: 109-115.
 - 3. Nakai K, Kanehisa M. 1991. Expert systems for predicting protein localization sites in Gram-negative bacteria. Proteins: Structure, Function, and Genetics 11:95-110.
 - 4. Nielsen H, Engelbrecht J, Brunak S and von Heijne G. 1997.

 Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering 10, 1-6.
- 5. Claros MG and G von Heijne. (1994). TopPred II: an improved
 software for membrane protein structure predictions. Comput. Appl. Biosci. 10: 685-686.
 - 6. Borodovsky M, Rudd KE, and EV Koonin. (1994). Intrinsic and extrinsic approaches for detecting genes in a bacterial genome. Nucleic Acids Res. 22:4756-4767.
- Struvye M, Moons M, Tommassen J. 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein J. Mol. Biol. 218:141-148.
- Aduse-Opoku J, Slaney JM, Rangarajan M, Muir J, Young KA, Curtis MA. 1997. The Tla receptor protein of *Porphyromonas gingivalis* W50: a homolog of the RI precursor (PrpRI) is an outer membrane receptor required for growth on low levels of hemin. J. Bacteriol. 179:4778-4788.
- 9. Needleman SB, Munsch CD. 1970. Ageneral method applicable to the search of similarity in the amino acid sequence of two proteins.
 30 J. Molec. Biol. 48: 443-453.

10

25

35

CLAIMS:-

1. An isolated antigenic *Porphorymonas gingivalis* polypeptide, the polypeptide comprising;

an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or an amino acid sequence at least 85%, preferably at least 95%, identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEO. ID. NO. 531 and SEQ. ID. NO. 532.

- 2. A polypeptide as claimed in claim 1 in which the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532.
- 3. A polypeptide as claimed in claim 1 in which the polypeptide comprises an amino acid sequence at least 85%, preferably at least 95%, identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532.
 - 4. A polypeptide as claimed in claim 1 in which the polypeptide comprises at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532.
- 5. A polypeptide as claimed in claim 1 in which the polypeptide 30 comprises;

an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532; or an amino acid sequence at least 85%, preferably at least 95%, identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532; or

15

at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532.

- 5 6. A polypeptide as claimed in claim 1 in which the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532.
 - 7. A polypeptide as claimed in claim 1 in which the polypeptide comprises an amino acid sequence at least 85%, preferably at least 95%, identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532.
 - 8. A polypeptide as claimed in claim 1 in which the polypeptide comprises at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532.
 - 9. A polypeptide as claimed in claim 6 in which the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386, SEQ. ID. NO. 424, SEQ. ID. NO. 425, SEQ. ID. NO. 434,
- 20 SEQ. ID. NO. 447, SEQ. ID. NO. 458, SEQ. ID. NO. 475, SEQ. ID. NO. 498, SEQ. ID. NO. 499, SEQ. ID. NO. 500, SEQ. ID. NO. 501, SEQ. ID. NO. 387,
 - SEQ. ID. NO. 400, SEQ. ID. NO. 411, SEQ. ID. NO. 419, SEQ. ID. NO. 420,
 - SEQ. ID. NO. 427, SEQ. ID. NO. 429, SEQ. ID. NO. 433, SEQ. ID. NO. 437,
 - SEQ. ID. NO. 438, SEQ. ID. NO. 443, SEQ. ID. NO. 444, SEQ. ID. NO. 448,
- 25 SEQ. ID. NO. 449, SEQ. ID. NO. 452, SEQ. ID. NO. 455, SEQ. ID. NO. 457,
 - SEQ. ID. NO. 459, SEQ. ID. NO. 461, SEQ. ID. NO. 462, SEQ. ID. NO. 463,
 - SEQ. ID. NO. 467, SEQ. ID. NO. 468, SEQ. ID. NO. 469, SEQ. ID. NO. 482,
 - SEQ. ID. NO. 484, SEQ. ID. NO. 485, SEQ. ID. NO. 494, SEQ. ID. NO. 508, SEQ. ID. NO. 509, SEQ. ID. NO. 510, SEQ. ID. NO. 520, SEQ. ID. NO. 521,
- 30 SEQ. ID. NO. 522, SEQ. ID. NO. 525, SEQ. ID. NO. 526, SEQ. ID. NO. 528,
 - SEQ. ID. NO. 389, SEQ. ID. NO. 390 and SEQ. ID. NO. 391.
 - 10. An isolated antigenic *Porphorymonas gingivalis* polypeptide, the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532
- less the leader sequence set out in Table 3.

- 11. An isolated DNA molecule, the DNA molecule comprising a nucleotide sequence which encodes the polypeptide as claimed in any one of claims 1 to 10 or a sequence which hybridises thereto under conditions of high stringency.
- 5 12. An isolated DNA molecule as claimed in claim 11 in which the DNA molecule comprises a nucleotide sequence selected from the group consisting of SEQ. ID. NO. 1 to SEQ. ID. NO. 264, SEQ. ID. NO. 529 and SEQ. ID. NO. 530.
- 13. A recombinant expression vector comprising the DNA molecule as claimed in claim 11 or claim 12 operably linked to a transcription regulatory element.
 - 14. A cell comprising the recombinant expression vector as claimed in claim 13.
- 15. A method for producing a *P. gingivalis* polypeptide comprising culturing the cell as claimed in claim 14 under conditions that permit expression of the polypeptide.
 - 16. A composition for use in raising an immune response directed against *P. gingivalis* in a subject, the composition comprising an effective amount of at least one polypeptide as claimed in any one of claims 1 to 10 and a pharmaceutically acceptable carrier.
 - 17. A composition as claimed in claim 16 in which the composition further comprises at least one DNA molecule as claimed in claim 11 or claim 12.
- 18. A composition as claimed in claim 16 or claim 17 in which the pharmaceutically acceptable carrier is an adjuvant.
 - 19. A method of treating a subject for *P. gingivalis* infection comprising administering to the subject a composition as claimed in any one of claims 16 or claim 18 such that treatment of *P. gingivalis* infection occurs.
 - 20. A method as claimed in claim 19, wherein the treatment is a prophylactic treatment.
 - 21. A method as claimed in claim 19, wherein the treatment is a therapeutic treatment.
 - 22. A composition for use in raising an immune response directed against *P. gingivalis* in a subject, the composition comprising an effective amount of at least one DNA molecule as claimed in claim 11 or claim 12 and a pharmaceutically acceptable carrier.

30

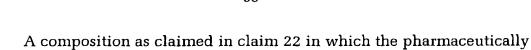
23.

15

25

30

acceptable carrier is an adjuvant.



- 24. A method of treating a subject for *P. gingivalis* infection comprising administering to the subject a composition as claimed in claim 22 or claim 23 such that treatment of *P. gingivalis* infection occurs.
- 25. A method as claimed in claim 24, wherein the treatment is a prophylactic treatment.
- 26. A method as claimed in claim 24, wherein the treatment is a therapeutic treatment.
- 10 27. An antibody raised against a polypeptide as claimed in any one of claims 1 to 10.
 - 28. An antibody as claimed in claim 27 in which the antibody is polyclonal.
 - 29. An antibody as claimed in claim 27 in which the antibody is monoclonal.
 - 30. A composition comprising at least one antibody as claimed in any one of claims 27 to 29.
 - 31. A composition as claimed in claim 30 in which the composition adapted for oral use.
- 20 32. A nucleotide probe comprising at least 18 nucleotides and having a contiguous sequence of at least 18 nucleotides identical to a contiguous nucleotide sequence selected from the group consisting of SEQ. ID. NO. 1 to SEQ. ID. NO. 121, SEQ. ID. NO. 529 and sequences complementary thereto.
 - 33. A nucleotide probe as claimed in claim 32 in which the probe further comprises a detectable label.
 - 34. A method for detecting the presence of *P. gingivalis* nucleic acid in a sample comprising:
 - (a) contacting a sample with the nucleotide probe as claimed in claim 32 or claim 33 under conditions in which a hybrid can form between the probe and a *P. gingivalis* nucleic acid in the sample; and (b) detecting the hybrid formed in step (a), wherein detection of a hybrid indicates the presence of a *P. gingivalis* nucleic acid in the sample.

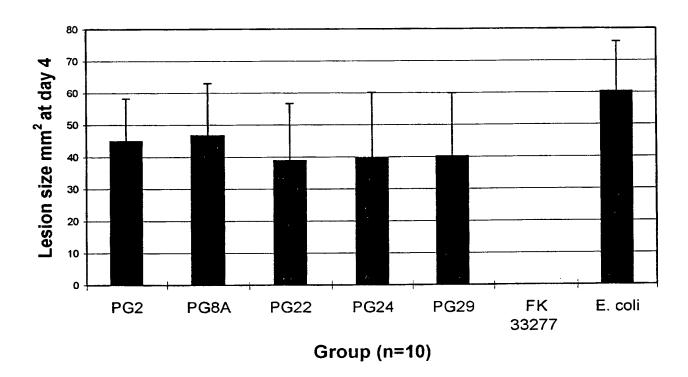


Figure 1

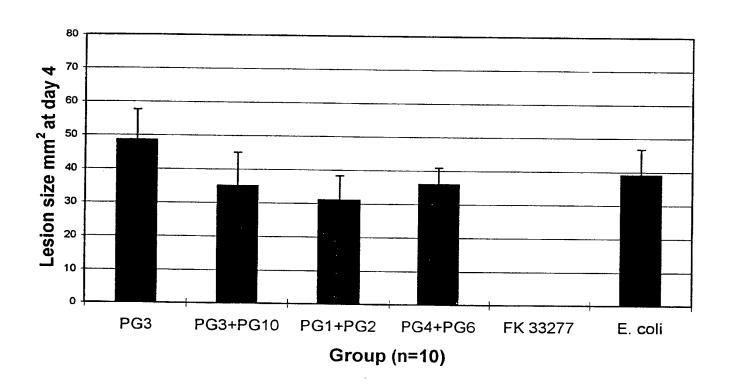


Figure 2



| | (2) INFORMATION FOR SEQ ID NO:1 | |
|------------|---|-------------------------------------|
| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1362 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 10 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 4.5 | (iv) ANTI-SENSE: NO | |
| 15 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 20 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11362 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1 | |
| 25 | TTCTGTGTCA TGGCAAAAGT TATAAAAACA AAAAAAGGCC TTGCACTTAA TCTGAAAGGA AAACCGCTGC CCGAGATGCT GGCCGAACCG GCCCAAAGTC CTACTTACGC GGTCGTGCCC GACGATTTTG AAGGTGTTAT CCCCAAGGTG ACGGCTCGTC CGGGGGATAA GGTGCGTGCC GGCTCAGCAC TGATGCACCA CAAGGCATAT CCGGAGATGA AGTTTACAAG TCCGGTTAGC GGCGAAGTGA TCGCGGTGAA TCGCGGTGCC AAGCGCAAGG TGTTGAGCAT CGAGGTGAAA | 60 120 180 240 300 |
| 30 | CCGGACGGAC TGAACGAATA CGAGTCATTC CCTGTCGGGG ATCCGTCTGC CCTCTTGCC GAACAGATCA AGGAGCTTTT ACTGTCGAGC GGTATGTGGG GTTTTATTAA GCAACGTCCT TACGACATAG TGGCTACACC GGATATAGCT CCACGCGACA TTTATATTAC TGCCAACTTT ACTGCACCAT TGGCTCCGGA CTTCGATTTC ATCGTTCGAG GAGAAGAACG CGCCCTGCAG ACTGCCATCG ATGCCTTGGC CAAACTCACG ACAGGAAAGG TGTATGTGGG CCTGAAGCCG | 360 420 480 540 600 |
| 35 | GGTTCATCTC TGGGCTTGCA CAATGCAGAA ATCGTAGAAG TACACGGACC TCATCCGGCA GGTAACGTGG GCGTGCTGAT CAATCATACG AAGCCAATCA ATCGGGGCGA AACGGTGTGG ACGCTCAAGG CTACCGACCT GATCGTGATC GGACGTTTCC TGCTTACCGG CAAAGCCGAT TTTACCAGAA TGATTGCCAT GACCGCTCCA GACGCTGCAG CTCACGGATA CGTCCGTATT ATGCCGGGTT GCAATGTCTT TGCTTCCTTC CCCGGCCGAC TGACAATAAA GGAATCTCAC | 660 720 780 840 900 |
| 40 | GAGCGTGTGA TCGATGGCAA TGTGCTGACC GGTAAGAAGC TCTGCGAGAA GGAGCCTTTC CTGTCAGCCC GGTGTGACCA GATCACGGTG ATCCCCGAAG GCGACGATGT GGACGAACTC TTCGGGTGGG CTGCACCCCG TCTCGATCAG TACAGCATGA GCAGAGCTTA TTTCTCTTGG TTGCAGGGGA AAAACAAAGA GTACGTACTC GATGCCCGGA TCAAGGGTGG CGAACGTGCT | 960 1020 1080 1140 1200 |
| 45 | CTCAAGGCTA TTATAGCATT CGACATCGAC AAGATGGAGG ACTTAGGCAT ATATGAAGTG GCTCCGGAGG ACTTTGCCAC TTGCGAATTT GTGGATACAT CCAAGATCGA GCTGCAGCGT | 1260 1260 1320 1362 |
| 50 | (2) INFORMATION FOR SEQ ID NO:2 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 603 base pairs (B) TYPE: nucleic acid | |
| 55 | (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 60 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 65 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 03 | .(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1603 | |
| 70 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2 | |
| <i>7</i> 5 | GAACTAAGCA AGTGTTACAT GGATAAAGTG AGCTATGCTC TGGGATTGAG CATCGGTAAT AATTTCAAGT CTTCGGGCAT CGACAGCGTC GTTATGGATG ATTTCATGCA AGGTCTGTCT GATGTACTGG AAGAAAAAGC CCCTCAGGTC TCGTATGACG AGGCCAAGCG CGAAATAGAG GCGTATTTCA TGGATTTGCA GCAGAAGGCT GTCAAACTGA ACAAAGAGC CGGAGAAGAA | 60 120 180 240 |

WO 99/29870 PCT/AU98/01023

| 5 | TTCCTCAAGA TAAATGCACA CAAGGAAGGT GTGACGACCT TACCGAGCGG CTTGCAATAC GAAGTCATTA AGATGGAGAG GGGCCGAAA CCCACCCTTT CGGACACGGT AACCTGTCAT TATCACGGTA CGCTCATCAA CGGTATCGTT TTCGATAGCT CTATGGACAG GGGAGAACCG GCCAGTTTCC CTCTAAGAGG AGTTATAGCC GGCTGGACGA AGATTCTTCA ATTAATGCCT GTAGGATCAA ACTGGAACAT TCAAACCGGG TAGTACGCTC ATTTTTATAA TCGAATTATT GAGTATCAAC AAA | 300 360 420 480 540 600 |
|----|---|--|
| 10 | (2) INFORMATION FOR SEQ ID NO:3 | |
| 15 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 837 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) HOLECULE TYPE: DHA (genomic) | |
| 20 | (iii) HYPOTHETICAL: NO | |
| | (iv) AMTI-SENSE: NO | |
| 25 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROHONAS GINGIVALIS | |
| 30 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1837</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3 | |
| 35 | CAAAAAAACA AACGAAAGAT GAAAAAAGCA TTACTTATTG GTGCTGCTCT TTTGGGAGCA GTCAGTTTTG CAAGTGCTCA GTCTTTGAGC ACAATCAAAG TACAGAACAA TTCAGTACAG CAACCTCGTG AGGAAGCCAC TATTCAGGTT TGTGGAGAAT TGGCAGAGCA AGTTGACTGC ATTGGACAG GTAATTCTGC AATCATAGCC GCTGCAGCGA AATTTGAAAG CGATGATCTC GAAAGCTATG TTGGCTGGGA GATCATGAGT GTTGATTTCT TCCCTGGATA TAAAGCGTGC | 60 120 180 240 300 |
| 40 | AAGTACACAT CTGCAGTCTG GGCTGATGAT ATGACCATCT TGGGCCAATC AGAAGATAGT GATCCCGAAA TGCAGACTATT CAACAATCTT GCTCTCAAGA CTAGTGTCAA GATTGAAGCC GGCAAGAATT ACATAGTTGG TTATATTGCT AATACCGCAG GTGGACATCC TATCGGATGT GATCAGGGCC CTGCCGTTGA TGGTTATGGA GATTTGGTTT CTATATCAGA AGATGGTGGT GCTACTTTCC CTCCGTTCGA ATCTCTTCAT CAAGCAGTTC CTACCTTAAA TTACAACATC | 360 420 480 540 600 |
| 45 | TATGTCGTTG TTCATTTGAA GAAGGGTGAA GGTGTTGAGG CTGTTCTTAC CAACGACAAG GCTAATGCTT ATGTTCAGAA TGGCGTTATC TATGTAGCCG GAGCTAATGG TCGTCAGGTA TCTCTGTTCG ACATGAACGG TAAGGTTGTT TATACCGGCG TTAGCGAAAC GATTGCAGCT CCTCAGAAGG GCATCAAGCT GGCTATC | 660 720 780 837 |
| 50 | (2) INFORMATION FOR SEQ ID NO:4 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 471 base pairs | |
| 55 | (B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 60 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: | |
| 65 | (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1471</pre> | |
| 70 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4 | |
| 75 | CGTGNATTCC TGCCGGAGAA AGCTCTCTAT ATCGGCTGCC GCGTGGAGAC GCAAGAGGGG CATGCCGTAG GTTTCGGACT GGATGACGC CCTGCGATGA AAGGCAAGGG CGATCTGGTC GGGAGCTATC TTCCCGGTGC TGCTCCGATG CCTTTTGTCC CGCTTTCTGA TATTCCGGCT CGTTCGATGG ATGCCAATTT TTATATCTAT TCTCGTATTT CACTGGGTTC GGGCACGCAG | 60 120 180 240 |

| 5 | GATGTTCTCC AGCATAGGAT GAAAGTGTAT CCCAACCCTG CCACGACGGA GCTGCATGTG GAAGCCCTGT CCGCATGGGT GGGCGAGCAG GCTGCGGTAT ATGATATGCG TGGTCGTCGG GTATCGGCTC GGACGGTGA TAGCGAGAAG CTGTGCATCG ACATTGCCTC ACTGCCCGTG GGCGTCTATA TGCTGCGCAT CGGCAGCTAC TCGGCCAAGT TCGAGAAGAG A | 300 360 420 471 |
|----|--|--|
| | (2) INFORMATION FOR SEQ ID NO:5 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1686 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 15 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HTPOTHETICAL: NO | |
| 20 | (iv) ANTI-SENSE: NO | |
| 20 | <pre>(vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS</pre> | |
| 25 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11686</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5 | |
| 30 | ACAAGAAATG TTTCACTCAT AAAAATGCCA AGAATTATGA AATTAAAAAT TGCACTCAGA CTGCTGCTGG CGACTTTTGC CATAGTTTTA TTTAGCCCTC TGGCCAAGGC CCAGATGGAT ATTGGTGGAG ACGATGTATT GATCGAGACG ATGTCCACCC TATCAGGATA TTCAGAGGAT | 60 120 180 |
| 35 | TTTTATTACA AGATGGCTGT GGCAGACAAT GGATGGATCT ATGTGATGTT GGATTTCTCT CGTATTTATT TTGATGATGT CAGGCTGTAT CGTTCCAAAG ACGTGGTGC TACTTACCAA AAGTTAGGGT CTTTGGGGTC TTTGGTGCCT TATGACTTCG ATGTCTCGCA TTGCGATTTT ATTGTAACGG GAAAGGATGA AGATGATATC AATGTTTGGA CAGTCATGAC AGCATTCGAA TATGTAGGTG GTACTATTGG CAATGGCGTT TTGCTGATGAC ATCGCCATGA TGCAGATATC AATAATACAG AGTGTGTGA CAAGAAGGAT TTCCCTAATA ATAGACTGAT GGGTGTAGCC | 240 300 360 420 480 540 |
| 40 | ATCGCCTCCA ACTACCGTGC GCCCTCTCCT TACGGTTTGG GGGGCGATCC TTTTGCTCTC GCTGTCGCCG TTAGTGGCTC CGGAAGCGAT CACAGCTTCT TGGACTATAT TTTTTCGTTA GATGGTGGAG TACACTTTGA GCAAAAGCGT ATTTACACAA GCCCCAAAA ACTGACTATC AATTACAGATAG ACCTTTCATT AGGCAGTACA TCTCCTTCTC TTGGATTTAA TACTTGGCCA CTAATGGAG TCGTATTCGA AATGAATAAG AACCTTGATG GCTTCGACAT TGGTTTCATT | 600 660 720 780 840 |
| 45 | TCCAACTTTG TGGACTATGA TCCCCGCTAT GCGTGGTCTG AACCGATAAT AATAGAAGAA GACTGTGGAT GGACTGATTT TAATCCTTTG GGAGCACTAA GTATAGAGAT CCAAATGATG TTGGATGACA ATTCGGATAA TACCGTGGGT GGAGAACGCT CCCATAACTT CCTGATCACT TACCCCGGGCC ATTACGTATA TCCGAAGCAA TCTTTCAATT ATTCTCCCGG ACATACACCG ACAAAGAAAG ATCTGGTCTT TAAACACTGT ATAGGTATTC CGGCTTTGGC ATACGATAAG | 900 960 1020 1080 1140 |
| 50 | GAAGGCGATC GTTATCTGAC TACTTTTCAA GATCACAATC TAATGAGATA CAGATGGATC AAATACGATG ACATTAACTC TTTTTATGGT TGGAGTTGGC CATATGTATA TGCAAAAGAA GCTAAAGATA AAAAGAGGCG CCGTCCGCAA GTAGCACTCA ATCCTACCAA TGGAAAGGCT TGTTGGGTAT GGCATACTCG CAAGAGCCCA TATGATGAAA CCAAACCACA TCCTACTCCT GTAATTATTA AACATTTCCT ATGGTCCGAT ACGGAGTGGG TACATGCTCT GGACGTGGGG | 1200 1260 1320 1380 1440 |
| 55 | GACGTATTGC AGAAGGAGGG TAGCATGAAG CTCTACCCCA ATCCTGCCAA AGAATATGTT CTGATCAACC TACCCAAAGA AGGGGGGCAC GAGGCAGTCG TATACGACAT GCAGGGCCGA ATCGTGGAGA AAGTTTCATT TTCAGGGAAA GAATATAAGC TGAATGTGCA GTATCTGTCC AAAGGTACGT ACATGCTGAA AGTTGTAGCG GATACGGAGT ATTTCGTGGA AAAAATCATT GTAGAG | 1500 1560 1620 1680 1686 |
| 60 | | |
| | (2) INFORMATION FOR SEQ ID NO:6 | |
| 65 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1173 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 70 | (ii) MOLECULE TYPE: DMA (genomic) | |
| 70 | (iii) HYPOTHETICAL: MO | |
| | (iv) ANTI-SENSE: UNKNOWN | |
| 75 | (vi) ORIGINAL SOURCE: | |



```
(A) ORGANISM: PORTPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
 5
                  (B) LOCATION 1...1173
           (x1) SEQUENCE DESCRIPTION: SEQ ID NO:6
       CAAATGAAAC GATTACTCCC CTTTCTCCTT TTAGCAGGAC TCGTAGCCGT AGGAAACGTG
10
       TCTGCTCAGT CACCCCGAAT CCCTCAAGTG GATGTACACA CTCGCATCGC AAGAAATGCC
       CGTTATCGAC TGGACAAGAT CAGTGTCCCG GATTCTCGTC AGATATTCGA TTACTTCTAT
                                                                                 180
       AAAGAAGAAA CGATACCCAC TAAAATACAA ACGACCACAG GAGGTGCAAT TACAAGCATC
       GATTCGCTTT TCTATGAAGA CGACAGGTTG GTTCAGGTGC GCTATTTTGA CAATAACCTT
                                                                                 300
       GAATTAAAAC AAGCGGAGAA GTATGTATAC GACGGTTCTA AGCTGGTCCT TCGAGAAATT
                                                                                 360
15
       CGCAAGTCGC CGACAGACGA AACGCCAATA AAGAAAGTTA GCTATCACTA TCTCTGTGGC
                                                                                 420
       AGCGATATGC CTTTTGAGAT TACGACAGAG ATGAGCGATG GCTATTTTGA AAGCCATACG
                                                                                 480
       CTTAACTATC TGAATGGAAA GATTGCCCGA ATAGATATCA TGACTCAACA GAACCCATCG
                                                                                 540
      GCCGAATTGA TCGAAACGGG TAGAATGGTA TATGAGTTTG ATGCCAATAA TGATGCTGTA
                                                                                 600
       CTGCTTCGTG ACAGTGTATT TCTTCCTCTT CAAAACAAGT GGGTAGAAAT GTTTACTCAC
                                                                                 660
20
      CGTTATACAT ACGACAATAA GCATAATTGT ATTCGTTGGG AACAAGACGA ATTCGGCACC
                                                                                 720
      CTCACCCTTG CCAACAACTT CGAATACGAC ACCACTATCC CTCTGTCGTC TGTATTGTTC
                                                                                 780
       CCCACGCATG AGGAGTTCTT CCGTCCTCTT CTTCCCAATT TTATGAAGCA TATGCGTACG
                                                                                 840
       AAGCAAACGT ATTTCAATAA CTCCGGAGAA GGCTTGTCAG AGGTATGCGA TTACAACTAC
                                                                                 900
      TTCTATACCG ATATGCAGGG TAATGCACTG ACCGATGTTG CCGTGAACGA ATCGATCAAG
                                                                                 960
25
       ATTTATCCTC GTCCTGCCAC GGATTTTCTG CGTATAGAAG GTTCGCAACT GCTTCGCCTT
                                                                                1020
      TCGCTATTCG ACATGAACGG GAAGCTCATC AGAGCTACCG AATTGACAGG CGATTTGGCC
                                                                                1080
       ATTATCGGAG TTGCATCTCT TCCGAGAGGC ACTTACATCG CAGAAATAAC TGCTGCAAAC
                                                                                1140
       AGCAAAACCA TACGTGCAAA AGTATCGCTC AGA
                                                                                1173
30
       (2) INFORMATION FOR SEQ ID NO:7
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1284 base pairs
35
                  (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
40
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
45
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
50
                 (B) LOCATION 1...1284
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7
      AAGAGGAATC CCTTACCTTT GACGGCATCA AATCGAAAGA TTTTTATAAA AATGAGACAG
55
       CATTTATCTC TATTTCCTTT TATCTTGTTT CTGCTTCTTG CCTTCTCTTA TGTCGGTTGC
                                                                                 120
       AGAACAGTCC GACAAACACC TAAGCAGTCG GAACGGTACG TCGTAGTCCT GTCTTTGGAC
                                                                                 180
       GGCTTCCGAC CGGACTATAC CGATCGGGCA CGTACACCGG CGTTGGATCG GATGGCACAG
                                                                                 240
      GAGGGATTGA GCGGGTCGCT CCAACCATGC TTCCCCTCGC TTACATTTCC CAATCATTAC
                                                                                 300
       AGCATGGCTA CGGGGCTTTA CCCCGATCAT CACGGTATCG TAGCCAATGA GTTTGTGGAT
                                                                                 360
60
       TCGCTACTGG GCATCTTTCG TATATCCGAC CGAAAAGCCG TGGAGACCCC CGGATTTTGG
                                                                                 420
      GGCGGCGAGC CGGTTTGGAA TACGGCCGCA CGCCAAGGCA TCCGTACCGG TGTCTACTTT
TGGGTAGGAT CCGAAACGGC TGTGAACGGA AATCGCCGT GGCGGTGGAA AAAATTCTCC
TCCACCGTTC CGTTTCGTGA CCGTGCCGAC TCCGTCATCG CGTGGCTCGG ACTGCCCGAA
                                                                                 480
                                                                                 540
                                                                                 600
      AAGGAGCGAC CGCGCTTGCT CATGTGGTAC ATCGAGGAGC CGGATATGAT CGGACACAGC
                                                                                 660
65
       CAAACGCCCG AAAGCCCGCT GACACTGGCA ATGGTAGAGC GGTTGGACAG TGTGGTCGGC
                                                                                 720
       TATTTCCGCA AGCGGTTGGA CTCTCTGCCC ATAGCCGCAC AGACCGACTT CATCATAGTA
                                                                                 780
       TCCGATCACG GTATGGCCAC GTACGAAAAT GAGAAATGTG TCAATCTGTC GCATTATCTG
                                                                                 840
      CCTGCGGACA GTTTCCTCTA CATGGCCACC GGGGCCTTCA CCCACTTGTA CCCGAAGCCC
                                                                                 900
      TCCTATACCG AGCGAGCCTA TGAGATCCTG CGGGCCATTC CACATATATC GGTTTACCGC
                                                                                 960
70
      AAGGGGGAGG TGCCCAAGCG TTTGCGCTGT GGCACCAATC CTCGTTTGGG CGAACTGGTC
                                                                                1020
      GTGATTCCGG ACATAGGCTC CACCGTCTTT TTCGCAATAA ATGAAGACGT TCGTCCGGGA
                                                                                1080
      GCGGCACATG GCTATGACAA CCAAGCACCG GAAATGCGGG CTTTACTCCG GGCTGTCGGA
                                                                               1140
       CCCGATTTCC GTCCGGCAG TAGGGTGGAA AACCTGCCGA ATATCACCAT CTATCCGCTC
                                                                               1200
      ATATGCAGGC TGTTGGGTAT AGAGCCTGCA CCCAACGATG CGGACGAAAC GTTGCTGAAC
                                                                               1260
75
      GGCCTGATCC GAGACAAACG ACCA
                                                                               1284
```

```
(2) INFORMATION FOR SEQ ID NO:8
  5
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 846 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
 10
           (ii) NOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
15
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
20
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...846
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8
25
      CTTTCCCGTG GAGITTTTCC CCTGATGTAT GGCCGAAGAG GAAGCATTCG TGCCTCTTCA
                                                                             60
      GGGCATAGGG ACAAAATTTT TAAGAATACA ATTATCAGAT TTATCACAAT GAAAGTAGGT
                                                                            120
      TTGTTCATCC CCTGTTATGT CAATGCAGTG TATCCGGAAG TGGGTATCGC CACGTACAAA
                                                                            180
      CTGCTGAAGA GTTTGGACAT AGATGTCGAC TACCCGATGG ATCAGACATG TTGCGGCCAG
                                                                            240
30
      CCTATGGCCA ATGCCGGATT CGAACAGAAA GCTCAAAAGC TGGCTTTGCG ATTCGAAGAG
                                                                            300
      360
      GAAAACTATG ATCATATCCT CAGACCGACA GGACATGTCT GCAAGTCGGC AGCCAAGGTT
                                                                            420
       CGGGATATAT GCGAGTTCTT GCACGATGAC CTGAAGATCA CCAGCCTCCC CTCCCGATTC
                                                                            480
      GCCCATAAGG TGAGCCTGCA CAACAGTTGC CACGGTGTGC GCGAACTGCA TCTGTCCACC
                                                                            540
35
      CCCAGTGAAG TGCACCGACC GTACCACAAC AAGGTGCGCC GGCTATTGGA GATGGTGCAG
                                                                            600
      GGCATAGAGG TATTCGAGCC GAAGCGAATA GACGAATGCT GCGGTTTCGG CGGTATGTAC
                                                                            660
      TCGGTGGAGG AGCCGGAGGT ATCCACCTGT ATGGGGCATG ACAAGGTGCT GGATCACATA
                                                                            720
      TCCACAGGTG CGGAGTACAT CACAGGGCCG GACAGCTCGT GCCTCATGCA TATGCAGGGA
                                                                            780
      GTGATAGACA GAGAGAAATT GCCGATCAAG ACAATTCATG CAGTAGAAAT TTTAGCAGCA
                                                                            840
40
      AACTTA
                                                                            846
      (2) INFORMATION FOR SEQ ID NO: 9
45
            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 753 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
50
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
55
          (iv) AUTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
60
          (ix) FEATURE:
                (A) MAME/KEY: misc_feature
                (B) LOCATION 1...753
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9
65
      CCTCTGAAAA AACGAATGGA TATTGTAAGT ATGGCCGATA AAGCTCTTGT AGTGGAGATG
                                                                            60
      AGAGATGTGA CGCTCTGTCA GGAGGAAAAC GTCATTTTTC AAAATTTGAA TCTGACCCTT
                                                                           120
      TCCGCCGGAG ACTTCGTCTA TCTGATAGGC TCAGTGGGAT CGGGGAAGAG CACTTTGCTG
                                                                           180
      AAGGCTTTGT ATGCTGAGGT GCCTATCTCT GCCGGTTATG CCCGCGTGAT AGATTATGAT
                                                                           240
70
      CTGGCAAAGT TGAAACGGAA GCAGTTGCCC TATCTGCGCA GGAATTTGGG CATTGTGTTT
                                                                           300
      CAGGATTTCC AGTTGCTGAA CGGACGTACT GTTGCGGAGA ATTTGGATTT CGTTTTGCGA
                                                                           360
      GCTACGGACT GGAAAAACCG AGCCGATCGC GAGCAGCGTA TCGAGGAGGT TTTGACCCGT
                                                                           420
      GTGGGAATGT CTCGGAAGGC TTATAAGAGA CCGCACGAAC TGTCCGGAGG GGAGCAACAA
                                                                           480
      CGTGTGGGTA TAGCCAGAGC TTTGCTGGCG AAGCCTGCGT TGATCCTGGC CGACGAACCC
                                                                           540
75
      ACAGGCAACC TCGATTCGGT GACCGGATTG CAGATCGCTT CTCTGCTCTA CGAAATCAGT
                                                                           600
```



| | AAGCAGGGCA CTGCAGTACT TATGAGCACG CACAACAGCA GCCTGCTGTC GCATCTGCCG GCACGGACAT TGGCCGTTCG TAAGAATGGC GATGCCTCCT CTTTGGTCGA GCTGAGTGCA GATGCTGTTT CAAGAAAAAA TACGGAAATA GAT | 660 720 753 |
|------------|---|--|
| 5 | (2) INFORMATION FOR SEQ ID NO:10 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 714 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 15 | (ii) HOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| 25 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1714 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10 | |
| 30 | ACCAGGCATT GTCCGGCTTG TCGCTCTTCC TTTCACCTCA TAAAAACAAG TAAAACAATG ATTGAAATCA GCAACCTCAC CAAGGTTTTC AGAACAGAG AAATAGAGAC GGTAGCCCTC GATGGCGTAT CGCTCAAAGT GGACAAAGGC GAATTTATCG CCATAATGGG GCCTTCGGGA TGCGGTAAGT CCACTCTGCT CAATATCCTC GGCCTTCTCG ACAATCCCAC TTCCGGTATC TACAAGCTCG ATGGGGCAGA AGTGGGCAAC CTCCGGGAAA AAGACAGGAC TGCCGTCCGT | 60 120 180 240 300 |
| 35 | AAGGGCAATA TCGGCTTCGT ATTCCAGAGC TTCAACCTCA TCGAAGAGT GACGGTAAGC GAGAACGTGG AGTTGCCGCT CGTCTATCTG GGTGTGAAGG CCTCCGAGCG GAAAGAGCGA GTGAGGGAGG CACTGCGCAA GATGAGCATC AGCCACCGGG CCGGCCACTT CCCCAAATCAG CTCCCGGAG GACAACAGCA GCGCGTGGCT ATCGCCCGTG CCGTGTGGC CAATCCGAAG CTCATCCTCG CCGATGAACC CACGGGTAAC CTCGACTCCA AAAACGGAGC CGATGTCATG | 360 420 480 540 600 |
| 40 | GAACTGCTCA GAGGTCTCAA TCGCGAAGGT GCAACCATCG TCATGGTGAC GCACTCCGAG CACGATGCAC GTAGTGCCGG CCGCATCATC AATCTGTTCG ACGGTAAGAT TCGC | 660 714 |
| | (2) INFORMATION FOR SEQ ID NO:11 | |
| 4 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1812 base pairs (B) TYPE: nucleic acid (C) STRANDEDMESS: double (D) TOPOLOGY: circular | |
| 50 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 55 | (iv) ANTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 60 | (ix) FEATURE: (A) NAHE/KEY: misc_feature (B) LOCATION 11812 | |
| 65 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11 | |
| | AGCACAGAAA CAAATAGCAA ATCTGAAATG AAAGAATTTT TCAAAATGTT TTTCGCCTCG ATCCTCGGGG TTATAACGGC AGGAATCATC TTGTTCTGTA TCTTTCTATT TATCTTTTTC GGCATCGTAG CCGGTATTGC CTCCAAGGCA ACGGGAGGAA CCATTCCGAA GATCGAAGCA | 60 120 |
| 70 | AACTCCATCC TACATATANA CAATTCTTCT TTCCCTGAGA TCGTATCGGC CAATCCCTGG AGCATGCTCA CAGGCAAAGA CGAGTCCGTA TCGCTCTCAC AGGCAGTCGA AGCCATCGGC CAAGCCAAAA ATAATCCCAA CATAACCGGT ATCTTCCTCG ATCTGGACAA CCTTTCCGTC GGTATGGCAT CGGCAGAGGA ATTGCGTCGC GCGTTGCAGG ATTTCAAGAT GTCGGGCAAG TTCGTCGTAT CCTATGCCGA CAGATACACC CAAAAGGGTT ACTACCTCTC CAGTATTGCA | 180 240 300 360 420 480 |
| 7 5 | GACAAACTCT ACCTCAATCC GAAAGGAATG TTGGGGCTTA TCGGGATTGC GACCCAAACA ATGTTCTACA AAGATGCCCT CGACAAATTC GGCGTGAAGA TGGAGATCTT CAAGGTAGGC | 540 600 |



```
ACCTACAAGG CAGCCGTAGA GCCATTCATG CTCAACAGGA TGAGCGATGC CAATCGCGAA
                                                                             660
                                                                              720
      CAAATCACCA CATACATAAA CGGGCTTTGG GACAAGATCA CATCCGATAT TGCAGAGTCG
                                                                              780
      CGCAAGACGG CAATGGATTC CGTGAAAATG TTTGCCGACA AAGGCGAAAT GTTCGGTCTT
      GCCGAGAAAG CGGTGGAGAT GAAGCTCGTG GATGAGCTGG CTTACCGTAC CGATGTGGAG
                                                                             840
 5
      AAAGAACTCA AAAAGATGTC CCAACGCGGA GAGAAAGATG AACTTCGGTT CGTATCGCTT
                                                                             900
      TCTCAGGTTC TGGCCAATGG CCCGATGAAC AAAACGAAAG GCAGTCGGAT CGCCGTTCTC
                                                                             960
      TTTGCCGAAG GTGAAATAAC GGAAGAAATA ATAAAGAAGC CGTTCGACAC TGACGGTAGC
                                                                             1020
                                                                             1080
      TCCATCACAC AAGAACTCGC CAAAGAAATC AAGGCAGCAG CCGATGACGA TGATATCAAA
      GCCGTAGTAC TTCGTGTCAA TTCTCCGGGA GGTAGTGCTT TCACTTCCGA ACAGATATGG
                                                                             1140
10
      AAGCAGGTAG CCGATCTCAA GGCCAAAAAG CCTATCGTGG TCTCCATGGG CGACGTAGCA
                                                                             1200
      GCCTCGGGCG GATACTACAT AGCCTGCGCA GCCAACAGTA TCGTGGCAGA GCATACGACT
                                                                             1260
      CTGACCGGCT CCATCGGCAT ATTCGGCATG TTCCCGAACT TCGCGGGCGT AGCCAAGAAG
                                                                             1320
      ATAGGAGTGA ATATGGACGT CGTACAGACA TCCAAGTATG CAGACTTGGG CAACACCTTC
                                                                             1380
      GCTCCGATGA CGGTCGAAGA TCGTGCCCTC ATCCAACGCT ACATAGAGCA GGGCTACGAC
                                                                             1440
15
      CTCTTCCTCA CTCGCGTATC GGAAGGCCGC AACCGCACCA AGGCACAGAT CGACAGCATC
                                                                            1500
      GCTCAAGGCC GTGTATGGCT CGGCGACAAA GCTCTTGCAC TCGGTTTGGT GGATGAGCTT
                                                                             1560
      GGAGGTTTGG ACACAGCTAT CAAACGGGCC GCGAAGCTGG CTCAGCTCGG TGGCAACTAC
                                                                             1620
      AGCATAGAGT ATGGCAAGAC CAAGCGCAAC TTCTTCGAAG AGTTGCTCTC CTCATCAGCA
                                                                            1680
      GCGGATATGA AGTCTGCCAT CCTGAGTACC ATTCTCTCCG ATCCGGAAAT AGAAGTTCTG
                                                                             1740
20
      CGCGAACTCC GCTCCATGCC GCCCGTCCT TCGGGCATAC AGGCACGTCT CCCCTATTAC
                                                                             1800
                                                                             1812
      TTCATGCCGT AC
      (2) INFORMATION FOR SEQ ID NO:12
25
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 972 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
30
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
35
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
40
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
                (B) LOCATION 1...972
45
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12
      CTAACGTTGT TTTGTTGCAA CTATTTCAAA CAGATGAGAG CAAACATTTG GCAGATACTT
      TCCGTTTCGG TTCTCTTTTT CTTCGGGACA GCGATCGGAC AGGCTCAGAG TCGAAACCGT
                                                                             120
      ACATACGAGG CTTATGTGAA ACAGTACGCC GACGAAGCTA TCCGACAGAT GAGCCGCTAC
                                                                             180
50
      AATATACCGG CAAGCATCAC CATAGCACAG GCTTTGGTGG AGACAGGAGC CGGAGCCAGT
                                                                             240
      ACACTGGCCA GCGTACACAA CAATCACTTC GGGATCAAAT GCCACAAATC GTGGACGGGC
                                                                              300
      AAGCGCACCT ATCGTACCGA CGATGCGCCG AACGAATGCT TCCGCAGCTA TTCGGCCGCT
                                                                             360
      CGCGAATCGT ATGAAGATCA TTCCCGATTT CTGCTCCAAC CACGCTATCG TCCCCTGTTC
                                                                             420
      AAACTCGACA GAGAAGACTA TCGGGGCTGG GCTACGGGGT TGCAACGCTG TGGCTATGCC
                                                                             480
55
      ACCAATCGGG GCTATGCCAA TCTGCTGATC AAGATGGTGG AGCTGTATGA GCTATATGCT
                                                                             540
      TTGGATCGCG AGAAGTACCC CTCATGGTTC CACAAGTCTT ACCCCGGGTC CAACAAAAA
                                                                             600
      TCCCATCAAA CGACCAAGCA GAAGCAGAGC GGACTCAAGC ACGAAGCTTA CTTCAGCTAC
                                                                             660
      GGACTGCTCT ACATCATAGC CAAGCAAGGC GATACCTTCG ATTCTTTGGC CGAAGAGTTC
                                                                             720
      GACATGAGAG CCTCCAAACT GGCCAAATAC AACGATGCTC CCGTGGATTT CCCGATCGAA
                                                                             780
60
      AAGGGCGATG TGATCTATCT GGAGAAAAAG CACGCATGCT CCATCTCCAA ACACACACAG
                                                                             840
      CACGTAGTGC GTGTGGGCGA TTCGATGCAC AGTATCTCCC AACGCTATGG CATCCGGATG
                                                                             900
      AAGAACCTCT ACAAGCTCAA CGACAAGGAT GGCGAATATA TACCCCAAGA GGGCGATATA
                                                                             960
      CTGCGCTTGC GC
                                                                             972
65
      (2) INFORMATION FOR SEQ ID NO:13
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1599 base pairs
70
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
75
```



| | (iii) HYPOTHETICAL: NO | |
|------------|---|---------------------------------|
| | (iv) ANTI-SENSE: NO | |
| 5 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 10 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11599 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13 | |
| 15 | AGGATCCCCG ACGAGCAGAC CGGACGTATC ATGGACGGAC GTCGATATTC GGATGGCCTC CATCAGGCTA TCGAAGCCAA AGAGCATGTG AAAGTAGAGG CTGCGACACA GACATTTGCA ACTATCACTT TGCAGAACTA TTTCCGCATG TATCATAAGC TGGCAGGGAT GACCGGTACT GCTGAAACTG AAGCGGGAGA GCTTTTGGGAC ATCTACAAAC TGGACGTTGT AGTTATTCCG | 60 120 180 240 |
| 20 | ACAAACAAGC CTATCGCCCG TAAGGATATG AATGATCGTA TCTATAAGAC GGCACGTGAA AAATATGCAG CAGTTATCGA AGAGATTGTA CTTCTTGTCG AAGAGGCCAG ACCTGTACTT GTCGGTACTA CTTCGGTGGA AATATCCGAA TTGTTGAGCC GTATGTTACG CTTGCGTGGC ATCCAACACA ATGTACTCAA TGCCAAATTG CATCAGAAGG AGGCCGAGAT TGTAGCTCAG GCCGGTCAGA AAGGAACTGT TACCATCGCA ACGAACATGG CCGGTCGTGG TACCGACATC | 300 360 420 480 540 |
| 25 | AAGCTCTCTG CCGAGGTTAA GAAAGCCGGG GGTTTGGCTA TCATTGGTAC GGAAAGGCAC GAATCCAGAC GAGTGGACAG ACAGCTTCGT GGTCGTTCCG GCCGTCAGGG TGATCCCGGT TCGTCCATAT TCTATGTTTC CCTTGAAGAT CATCTGATGC GCCTCTTTGC CACAGAAAAG ATTGCATCAT TGATLGATCG TTTAGGTTTC AAGGAAGGAG AAGTGCTCGA AAACAACATG | 600 660 720 780 |
| 3 0 | CTGAGTAAGT CCGTGGAGCG TGCTCAAAAG AAGGTGGAAG AGAACAACTT CGGTATCCGT AAACATCTGC TTGAGTACGA TGATGTAATG AATTCGCAGC GTGAAGTCAT TTATACCCGT CGCCGTCATG CTTTGATGGG AGAGCGTATC GGTATGGATG TACTCAATAC CATATACGAC GTATGTAAGG CTCTGATTGA CAATTATGCA GAAGCCAATG ATTTCGAAGG CTTCAAGGAA | 840 900 960 1020 |
| | GATCTGATGC GTGCACTCGC GATAGAATCT CCTATCACGC AAGAAATATT CAGAGGTAAG AAAGCAGAAG AGCTGACCGA TATGCTTTTC GATGAAGCTT ACAAGTCTTT CCAACGTAAG ATGGATCTGA TCGCAGAAGT GGCCCACCCT GTGGTTCATC AGGTATTCGA GACCCAAGCC GCCGTGTACG AGCGCATTCT AATCCCCATT ACGGATGGTA AACGTGTCTA TAACATAGGA | 1080 1140 1200 1260 |
| 35 | TGCAATTTGC GTGAAGCGGA TGAAACTCAA GGGAAAAGCA TCATCAAAGA ATTTGAGAAA GCTATCGTAC TGCATACTAT CGATGAGTCT TGGAAAGAAC ATCTGCGTGA GATGGACGAG CTTCGTAATT CCGTTCAGAA TGCCAGCTAC GAAAACAAAG ATCCACTACT TATCTATAAA CTCGAATCTT ACGAACTGTT CCGCAAGATG GTAGAAGCCA TGAACCGTAA GACCGTAGCG | 1320 1380 1440 1500 |
| 40 | ATCCTAATGC GTGCTCGGAT ACCGGTACCG GAGGCTCCTT CCCAAGAAGA GCTGGAACAC AGGCGGCAAA TAGAAATCCG ACATGCAACC CAACAACGT | 1560 1599 |
| | (2) INFORMATION FOR SEQ ID NO:14 | |
| 45 | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 2160 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double | |
| 50 | (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 55 | (iv) AHTI-SENSE: NO | |
| | <pre>(vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS</pre> | |
| 60 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12160</pre> | |
| 65 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14 | |
| | AAATCTTGCC GTGTTATTGG CCAGACGAGG CGATACGGAT GCTGCCCTGT CGGACTACGA CCGTGCCATC AAAGCCTATC CGGAGTTTGC CGATGCCTAT TTCAATCGCG GCCTGCTGTT GCTTTCGCGC GGAAAGGCCA AAGAAGGCAT CGCCGATCTG AGTCGGCAG GCGAATACGG | 60 120 180 |
| 70 | GCTCTACAAG GCGTACAACA TCATCAAACG AATGAGCACG AAGTCATGAT CTCCGTCAAT AACCTGACTG TCGATTCGG CACCCGTCTG CTCTTCGATC AGGTATCATT CGTCATCAAC AGGCGCGACC GTATCGCTCT TGTAGGGAAG AACGGTGCCG GCAAGAGTAC GCTGCTCAAG CTGATTGCCG GCATGGAAGA ACCGACATCC GGACACATAG CACGCCCCAA GGGGATCCGC | 240 300 360 420 |
| 75 | ATAGGCTATC TGCCGCAGGT GATGCGTTTG CAGGACGGAC ACACGGTTTA CGAAGAGGTC GAGCAGGCTT TCAACGATAT TCGCCAAATA GAGGAAGAGA TACGGCGTCT GTCCGATGAA ATGGCCGGAC GTACGGACTA CGAATCGGAT GACTATATCC GACTGATAGA GCATTATACG | 480 540 600 |

PCT/AU98/01023

9 / 490

```
AATATGAGCG AGACCCTCTC TCTCATGCAG CAGGGCAACT ATCATGCTGC GATCGAACAG
                                                                                     660
        ACATTGATCG GTCTGGGCTT CGGCCGAGAG GACTTCCACC GCCCCACAGC CGATTTCAGC
                                                                                     720
        GGAGGATGGC GTATGCGGAT AGAGCTGGCC AAACTTCTGC TCCAACGCCC CGAAGTTTTG
                                                                                    780
        CTGCTCGACG AGCCGACCAA TCACCTCGAC ATCGAATCCA TCGGCTGGCT GGAGCAGTTC
                                                                                    840
  5
        ATCGCCACCA ATGCAGGAGC CGTTATCCTG GTGTCGCACG ACAGGGCATT CATCGACAAT
                                                                                    900
        ACCACGACAC GCACTATCGA AATAGAACTG GGACATATAT ACGACTACAA GACCAACTAC
                                                                                    960
       AGCCACTATG TGGAGCTGCG CGAAGAGCGG CTGCGACAGC AGATGCGTGC CTACGAGAAT
                                                                                   1020
       CAGCAGAAGA TGATCCGCGA TACGGAGGAC TTCATCGAAC GATTCAGATA CAAGGCCACG
                                                                                   1080
       AAGTCCGTAC AGGTACAGAG CCGGATCAAA CAGTTGGAGA AAGTAGAGCG CGTGGAGATA
                                                                                   1140
10
       GACGAGCGG ATCGTTCGGC ATTTCACTTC CGCTTTATCC CGGCACAGCC TTCCGGCAGT
                                                                                   1200
       TATCCGCTAA TAGTGGATGA TTTGGCCAAG GCTTATGGCG ATCACCAGGT GTTTTCCGGA
                                                                                   1260
       GCTACATACA CCATCGAAAG AGGCGAAAAG GTGGCTTTCG TAGGCAAAAA CGGTGCCGGC
                                                                                   1320
       AAAAGTACCA TGGTCAAGTG TATCATGGGA GAGCTGACAG ACTACACCGG CAAGCTCGAA
                                                                                   1380
       CTGGGGCACA ACGTGCAGCT GGGCTACTTT GCCCAAAACG AAGCCCAAGA GCTAAGAGGG
                                                                                   1440
15
       GATCTCACGG TATTCGACAC GATAGACCGT GAGGCCGTGG GCGACATCCG TCTGCGCCTG
                                                                                   1500
       AACGATTTGC TCGGGGCTTT TCTCTTCGGG GGCGAAGCAT CGGAAAAGAA AGTAAGTGTC
                                                                                   1560
       CTGAGTGGAG GAGAACGAGC ACGATTGGCT ATTATCAGGC TTTTGCTACA GCCGGCTAAC
                                                                                   1620
       TTCCTTATTC TCGATGAGCC GACCAATCAC CTCGATATGC GCTCGAAGGA TGTACTGAAA GAGGCGATCA AGAACTTCGA TGGGACTGTC ATCGTAGTAT CTCACGACCG TGAGTTCCTC
                                                                                   1680
                                                                                   1740
20
       GATGGGCTTG TCAGCAAGGT GTATGAATTT GCAGATGGAC AGGTGAACGA ACACCTCGGA
GGTATATACG ACTATCTCCG GACCCGCCGT ATGCAGACGC TGACAGAGCT GGAGCGAACC
                                                                                   1800
                                                                                   1860
       ACTACGATCG AAACAAAAAC CACACGGGAG GCTATACCTG AAACGGAAGC CAAAGCGGAC
       TACCGTCGGC AAAAGGAGGT AGCCAAACAG CTGCGCACGT TGGAGCGAAC CGTAGCAACC
                                                                                   1980
       TGCGAGGAGC GGATCGGAAA ATTGGAGTCG GAATTACAGG CAATAGAGAT GCTACTGCAA
                                                                                   2040
25
       GATCCGAAAC ATGCGACTGA CGCGAATCTG TTCGAGCGAT ACGCCGGCAT GAAACAAGAA
                                                                                   2100
       CTCGAAAAGG CCATGGAGGA CTGGGAACAG GCTTCCGAAG CTTTATCCGA AGCCCAAGGA
                                                                                   2160
       (2) INFORMATION FOR SEQ ID NO:15
30
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1158 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
35
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
40
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1158
50
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15
       TATTTGAAGC TATTAATACT ACAAATAGCC TTGATGAATT TCTTAAAAAA AGAACCGTTT
                                                                                     60
       AAAATATTCT CTATGATTTA TCTGCTGTTA GATACAATAA CAAACCGTGC CGGTACAGAA
                                                                                    120
       CGCGCCGTGA TCAACTTGGC TAACAACCTG CATGCCAATG GTCATCGCGT ATCATTAGTC
                                                                                    180
55
       AGCGTTTGTA CAAAAGAAGG AGAGCCTTCC TTCCAAGTAG AAAAAGGAAT AGAAGTACAC
                                                                                    240
       CATCTCGGAA TTAGGCTTTA TGGCAATGCA TTAGCCCGCA AAACAGTATA TTTCAAGGCT
                                                                                    300
       TATCGAAGGA TAAAAGCCCT ATACAAGAAG CGTGAACCGG TTTTATTGAT AGGGACTAAT
                                                                                    360
       ATTTTTATCA ATACAATTTT GTCTCAGATC AGTAACAGAG GCAGAATATT TACGATCGGA
                                                                                    420
      TGCGAACATA TCTCTTATGA TATTGCCCGC CCTATTACAA AACGCATAAG GGGGTTTCTG
TATTCAGGGC TTGATGCCGT TGTAGCACTG ACAAAAAGAG ATCAGCAATC GTTCGAGGCA
                                                                                    480
60
                                                                                    540
      ATCTTACGTG GACGCTCTAA AGCATATGTC ATACCCAATC AAGTTTCATT TACTACAGTC
      CAAAGAGATG CTACTACTCA CAAACAAATG TTGGCGATTG GCAGGCTTAC CTACCAGAAG
                                                                                    660
       GGTTTTGAAT TCATGATAGA AGATGCATCA CGAGTGCTGC GAGAAAGGCC TGATTGGAAG
                                                                                    720
      CTTATCATAG TCGGAGATGG CGAAAATGAA TCGATGCTAC GTAAAGAAAT TGCATCTCGC
                                                                                    780
65
       AATATGGAGT CGCAAATAGA AATACATCCA TCTACACCGG AAATTCGCAA ATACTACGAA
                                                                                    840
      TCATCTGCTA TTTATCTAAT GACGTCCCGT TTCGAAGGAC TACCAATGGT ACTTCTCGAA
GCAGAAGCAT ATGCACTACC TATAATCTCA TACGATTGTC CGACCGGCCC GAGGGAACTG
                                                                                    900
                                                                                    960
      ATCGAAAACG GTCGCAATGG TTTCCTTGTG CCAATGGAAG CACATGAAGA CTTCGCGGAT
                                                                                   1020
      AAGTTACGCT TATTGATGGA TGATGAAACT CTTCGTAAGA AAATGGGACA AGAATCAGAG
                                                                                   1080
70
       TTGATGGTCA AATCCTACTC TCCGGCAAAT ATCTATGAAT GTTGGAAGAA ACTATTCGTC
                                                                                   1140
      GAAATCGGCT ACATGAAT
                                                                                  1158
       (2) INFORMATION FOR SEQ ID NO:16
```

```
(i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1965 base pairs
                 (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
 5
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
10
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
15
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...1965
20
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16
      ACAACGAAAG AAAACGAGAC AACAACAAAA AACGAATATA GAATTATGGG AAAAATCATT
      GGAATTGACT TAGGCACAAC GAACTCTTGT GTCTCTGTAT TGGAAGGTAA CGAACCTATC
                                                                             120
      GTTATTACAA ACAGTGAGGG CAAGCGCACA ACGCCCTCGG TAGTGGCTTT TGTGGATGGT
                                                                             180
25
      GGCGAGCGTA AGGTGGGCGA TCCGGCCAAG CGTCAGGCCA TCACCAATCC GACCAAGACG
                                                                             240
      ATATACTCTA TCAAACGCTT CATGGGCGAA ACTTACGATC AGGITTCCAG AGAAGTGGAG
                                                                             300
      AGAGTGCCAT TCAAGGTAGT ACGTGGGGAC AATAATACTC CGCGCGTAGA TATAGACGGT
      CGTCTCTATA CGCCGCAGGA AATTTCGGCC ATGATCCTTC AGAAGATGAA GAAGACGGCC
                                                                             420
      GAAGACTACC TCGGTCAGGA AGTAACGGAG GCCGTGATCA CTGTGCCCGC ATACTTCAAC
                                                                             480
30
      GACGCTCAAC GTCAGGCAAC GAAAGAAGCA GGAGAGATCG CCGGCCTGAA AGTTCGCCGT
                                                                             540
      ATTGTGAACG AGCCTACGGC AGCTTCTCTG GCCTACGGTC TGGACAAGTC CAATAAGGAT
                                                                             600
      ATGAAGATCG CTGTCTTCGA CTTGGGTGGC GGTACCTTCG ATATCTCTAT CTTGGAATTG
                                                                             660
      GGCGACGGCG TTTTCGAAGT GAAATCGACC AACGGTGATA CGCACCTCGG AGGAGACGAC
                                                                             720
      TTCGACCACG TGATCATTGA CTGGCTGGCA GAAGAGTTCA AGTCTCAGGA AGGTGTGGAT
                                                                             780
35
      CTTCGCCAGG ATCCTATGGC TATGCAGCGT CTGAAAGAAG CTGCCGAAAA AGCCAAGATA
                                                                             840
      GAGCTCTCCA GCACTTCATC TACGGAGATC AACCTCCCCT ATATCATGCC GGTGAACGGC
                                                                             900
      ATCCCCAAGC ACTTGGTGAT GACGCTTACA AGGGCTAAGT TCGAGCAGTT GGCCGATCGT
                                                                             960
      CTGATTCAGG CATGTGTGGC ACCCTGCGAA ACGGCCTTGA AAGATGCCGG TATGTCACGT
                                                                            1020
      GGCGATATCG ATGAAGTGAT TCTCGTAGGT GGTTCCACAC GTATTCCTGC TATTCAGGAG
                                                                            1080
40
      ATTGTGGAGA AGATCTTCGG TAAGGCTCCG TCCAAGGGTG TGAATCCCGA CGAAGTGGTA
                                                                            1140
      GCTGTGGGTG CCGCTATTCA AGGCGGTGTT CTGACCGGTG AGGTAAAGGA TGTCTTGCTG
                                                                            1200
      TTGGACGTTA CCCCCTTGTC GCTCGGTATC GAGACTATGG GAGGCGTGAT GACTCGCTTG
                                                                            1260
      ATCGATGCCA ATACCACTAT CCCGACGAAG AAGAGCGAAA TCTTTACCAC AGCAGTGGAC
                                                                            1320
      AATCAACCTT CGGTAGAGAT TCATGTACTT CAGGGTGAGC GTTCTTTGGC TAAGGACAAT
                                                                            1380
45
      AAGAGCATCG GCCGTTCAA CTTGGACGGT ATTGCTCCGG CGCCCCGTCA GACACCGCAG
                                                                            1440
      ATCGAAGTAA CGTTTGACAT CGATGCCAAC GGTATCCTGA ATGTAACGGC TCATGACAAA
                                                                             1500
      GCTACCGGCA AGAAGCAGAA TATCCGCATC GAAGCCTCCA GCGGTTTGTC CGATGATGAG
                                                                             1560
      ATCAAGCGCA TGAAGGAAGA GGCGCAGGCC AATGCCGAAG CAGATAAGAA AGAGAAAGAA
                                                                            1620
      CGTATCGACA AGATCAATCA GGCCGACAGC ATGATCTTCC AGACGGAAAA GCAGTTGAAG
                                                                            1680
50
      GAGTTGGGAG ACAAATTCCC GGCCGACAAG AAGGCTCCGA TCGATACCGC TCTCGACAAA
                                                                            1740
      CTGAAAGAAG CACACAAAGC ACAGGATGTA GCTGCTATCG ATACAGCCAT GGCCGAACTG
                                                                            1800
      CAAACCGCTC TTTCCGCAGC GGGCGAAGAG CTTTACAAGA ATGCCGGAGC AGCCCAAGGT
                                                                            1860
      GGCGCACAAC CCGGTCCGGA CTTCGGCGGT GCTCAAGGTC CCTCTGCCGG TGATCAGCCC
                                                                            1920
      TCTGACGACA AGAACGTCAC AGACGTAGAC TTCGAGGAAG TGAAG
                                                                            1965
55
      (2) INFORMATION FOR SEQ ID NO:17
           (i) SEQUENCE CHARACTERISTICS:
60
                 (A) LENGTH: 1401 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
65
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) AHTI-SEHSE: NO
70
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
75
                (A) NAME/KEY: misc_feature
```

(B) LOCATION 1...1401

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

| 5 | AAGTGGGCAC GTACAACTAC TTTAAGGATA TCGAACGCAA CAATCTATAA AACTATGCGC TACGACTAG CTATCATCGG TGGAGGGCCG GCCGGTTATA CGGCTGCCGA ACGTGCTGCC AAAGGTGGCC TGAAAACCCT CCTAATTGAG AAGAATGCTC TCGGTGGTGT ATGCCTCAAC GAAGGATGTA TACCGACCAA GACGCTACTC TACTCGGCCA AAGTGCTACA TCAAATTGCT ACGGCATCTA AATATGCAGT AAGTGGAACG GCCGATGGAC TTGACCTCGG CAAGGTGATT | 60 120 180 240 300 |
|------------|--|--|
| 10 | GCCAGAAAAG GTAAAATCAT TCGCAAGCTG ACTGCAGGCA TCCGTTCACG CCTGACAGAG GCCGGAGTAG AGATGGTGAC GGCAGAAGCT ACCGTAACGG GATGCGATGC | 360 420 480 540 |
| 15 | GGAATGGAGT TCGTCTTC CTTCAACGGT ATCGGTGG TGGAGTGATC GGAATGGAGT TCGTCTTC CTTCAACGGT ATCGGTAGA AAGTGCACGT GGTGGAGATG CTGCCGGAAA TACTCAACGG TATCGATCCC GAACATGCAG CTATGCTACG CGCTCACTAT GAAAAAGAAG GAATCAAATT CTACCTCGGG CACAAAGTAA CATCGGTTCG CAACGGAGCT GTTACGGTAG AATACGAAGG AGAAAGCAAA GAGATCGAAG GAGAACGTAT CCTGATGAGT GTGGGACGTC GCCCCGTGCT GCAAGGATTC GAGTCGCTCG GATTGGTGCT TGCCGGCAAA | 600 660 720 780 840 900 |
| 20 | GSTGTAAAGA CTAATGAGAG GATGCAAACT TCCCTGCCA ATGTCTATGC TGCAGTGAT ATTACAGGCT TCTCGCTTTT GGCACATACG GCTGTACGGG AAGCAGAGGT AGCAGTAGAT CAGATTTTGG GCAAAACAGA CGAAACGATG AGCTACCGTG CCGTACCAGG TGTGGTGTAC ACCAATCCCG AGGTCGCCG TGTGGGAGAG ACGGAAGAAT CGCTTCGCAA AGCAGGACGT GCCTACACTG TTCGTCGCCT TCCTATGGCC TTCTCCGGTC GATTTGTAGC AGAAAACGAA | 960 1020 1080 1140 |
| 25 | CAAGGCAATG GAGAGTGCAA ACTACTACTT GATGAAGAAA ACCGCTTGAT CGGAGCACAC CTCATTGGCA ATCCGGCCGG CGAACTCATC GTAACCGCT CCATGGCCAT CGAGACCCGC ATGACCGATC GACAAATCGA ACGAATCATA TTCCCTCATC CGACTGTAGG CGAAATCCTA AAAGAAACTC TCGCCGGAGG T | 1200 1260 1320 1380 1401 |
| 30 | (2) INFORMATION FOR SEQ ID NO:18 | |
| 35 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2835 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 4 0 | (ii) MOLECULE TYPE: DHA (genomic) | |
| | (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO | |
| 45 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| 50 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12835</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18 | |
| 55 | CCTAAGATAC TTATGGAATT GAAAAGATTT TTATCACTTG GTCTTCTGCT TGTGGGATTC ATTCCGATGA AGCTTTCTGC CCAACAGGCT CAGCCACTCC CTACAGATCC GGCTGTTCGT GTCGGTAAGT TGGACAACGG ATTGACTTAT TTCATCCGTC ACAACGAGAA CCCGAAAGAT CGTGCGGATT TCTTTATCGC ACAAAAGGTA GGTTCTATTC TTGAAGAAGA TAGCCAGTCC GGTTTGGCTC ACTTCTTGGA ACACATGGCT TTCAACGGTA CGAAGAACTT CCCCGGTAAG | 60 120 180 240 300 |
| 60 | AACTTGATCA ACTATCTCGA AACGATCGGT GTACGTTTCG GTCAGAACCT GAACGCTTCT ACCGGATTCG ACAAGACGGA ATATACGATA ATGGATGTGC CGACTACACG TCAGGGAATC ATCGACTCCT GCTTGCTTAT CCTGCATGAT TGGAGTAACA ATATTACCCT CGACGGGCAT GAGATCGACG AGGAGCGCG TGTGATCCAG GAAGAGTGGC GTGCTCGTCG CGATGCCAAC CTTCGTATGT TCGAGGCTAT ACTTGCCAAG GCTATGCCGG GTAATAAATA TGCAGAACGC | 360 420 480 540 |
| 65 | ATGCCCATCG GTCTGATGGA CGTCGTGCTC AACTTCAAGC ATGATGAGCT GCGCAACTAT TATAAGAAAT GGTATCGTCC CGACCTGCAA GGTCTGGTGA TCGTGGGAGA TATCGATGTG GACTATGTGG AGAACAAGAT CAAAGAACTC TTCAAGGACG TTCCTGCTCC CGTGAATCCA GCAGAGCGTA TCTATACGCC GGTAGAGGAC AACGATGAGC CTATCGTAGC CATTGCTACC | 600 660 720 780 840 |
| 70 | GATGCTGAGG CTACTACCAC GCAGCTCTCC ATCAGCTTCA AGAGCGACCC CACTCCTCAA GAAGTGCGAG GATCATATC CGGACTTGTG GAAGACTATA TGAAACAGGT GATCACTACA GCCGTGAATA AGCGTCCTTC CACAAGCCTA ACGCTCCTTT CCTCAGTGCA GCACAGGCTTC TCTCTAACTT CATGTACATC ACCAAGACTA AGGACGCATT CAATTTTGTT GCCACGGTTC GTGAGGGTGA AGCGAGAAA GCGATGAACG CATTGGTGC AGAGATAGAA AGCCTCCGTC AGTTCGGTAT CACCAAAGGC GAATACGATC GTGCACGCAC GAATGTGCTC | 900 960 1020 1080 1140 1200 |
| 75 | AAGCGATACG AGAATCAATA CAACGAAAGA GACAAGCGTA AGAACAATGC TTATGCCAAT GAATACTCCA CCTACTTCAC CGATGGCGGC TATATCCCGG GTATTGAGGT GGAATATCAG | 1260 1260 1320 |
| | | |

WO 99/29870 PCT/AU98/01023

| | ACGGTGAATG CTTTTGCTCC TCAGGTTCCT CTGGAAGCAT TCAATCAGGC TATTGCCCAA | 138 |
|------------|---|--------------|
| | ATGATCGATC CGGTGAAGAA TGCTGTCGTT ACCCTCACCG GTCCTTCAAA GGCTGAAGCC | 144 |
| | AAGATTCCGA GCGAAGCAGA CTTCCTCGCT GCTTTCAAAG CTGCTCGTCA GCAGAAAGTA | 150 |
| 5 | GAAGCCAAGA AAGACGAAGT CTCCGACCAA AAATTGATGG AGAAAGCTCC TAAGGCCGGA | 156 |
| , | AAGATCGTTT CCGAGAAGAA AGATCAGAAG TTCGGTACCA CGGAACTTAC CCTTAGCAAT | 1620 |
| | GGCATCAAAG TATACCTCAA GAAGACCGAT TTCAAATCAA ACGAAATCCT GATGAGTGCT CTCAGCCCGG GTGGTATCCT CTCCGGAAAG CATGCTCCCA ACCAATCTGT GATGAATTCG | 1680 |
| | TTCATGAACG TGGGTGGCTT GGGCAACTTC GATGCTATCC AGCTGGATAA GGTGCTGACA | 1740 |
| | GGTCGCTCTG CTTCCGTATC TCCCTCTTTG TCTCTGCTCA GTGAAGGTCT TTCGGGCAAA | 1800 1860 |
| 10 | ACGACTGTAG AAGATATGGA AACTTTCTTC CAGTTGATCT ATCTCCAAAT GACTGCTAAC | 1920 |
| | CGCAAGGATC CCGAAGCGTT CAAGGCCACA CAGGAAAAGT ТGTACAATAA СТТСААДАДТ | 1980 |
| | CAGGAAGCCA ACCCGATGGC TGCGCTTATG GACTCTATCC GTCATACCAT GTACGGCGAT | 2040 |
| | AATCCGATGA TGAAACCCAT GAAAGCTGCT GACGTGGAGA AAGTAAATTA CGATCAGGTA | 2100 |
| 15 | ATGGCTTTCT ACAATGAGCG ATTCGCTGAT GCCGGCGACT TTATGTTCTT CTTTATCGGT | 2160 |
| 13 | AATCTGGATG AAGCCAAGAT GAAGCCATTG ATCGAAACTT ATCTTGCTTC ATTGCCCAAC | 2220 |
| | CTCAAGCGTG GCGATAAGAT GAATAAGGCT CAGGTACCGG CTGCCCGTTC GGGAAAGATC GATTGCAAGT TCGAGAAGGA AATGGATACT CCTTCGACTA CTATATTCGA TGTCGTGTCC | 2280 |
| | GGAAATGTGG AATATACGCT CAAGAACAGT CTCCTGCTGG AAGTCTTCTC AGCCGTAATG | 2340 |
| | GATCAGGTGT ACACGGCTAC CGTTCGCGAG AAGGAAGGCG GTGCATACAG TGTGGCTGCA | 2400 |
| 20 | TTCGGCGGTC TCGAGCAATA TCCTCAGCCC AAGGCTCTGA TGCAGATCTA TTTCCCCACG | 2460 |
| | GATCCTGCTC GTGCCGAGGA AATGAATGCT ATCGTTTTTG CTGAGTTGGA GAAGCTTGCC | 2520 2580 |
| | AAGGAGGCC CCAATGTGGA ATACTTTAAG AAGACTATCG AAAACCTGAA TAAGCAGCAC | 2640 |
| | AAAGAAAGTC TGCGTGAGAA TCGTTTCTGG CTCGAAGCCA TGAAGGCGTC TTTCTTCGAA | 2700 |
| 25 | GGAAATGACT TCATCACAGA CTACGAATCC GTACTGAACG GTCTTACTCC TGCTGAATTG | 2760 |
| 25 | CAAAAGTTTG CEGCAGACCT CTTGAAGCAG CAGAATCGGG TTGTTGTCAT GATGGCTCCT | 2820 |
| | GTTGCAAAGG CTCAA | 2835 |
| | | |
| | (2) INFORMATION FOR SEQ ID NO:19 | |
| 30 | Total | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 2058 base pairs | |
| | (B) TYPE: nucleic acid | |
| 35 | (C) STRANDEDNESS: double | |
| 33 | (D) TOPOLOGY: circular | |
| | (ii) HOLECULE TYPE: DNA (genomic) | |
| | (12) Hobboods IIIB. BIN (Genomic) | |
| 40 | (iii) HYPOTHETICAL: NO | |
| 40 | (in) NUCL COVER NO | |
| | (iv) ANTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: | |
| | (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 4 5 | | |
| | (ix) FEATURE: | |
| | (A) NAME/KEY: misc_feature | |
| | (B) LOCATION 12058 | |
| 50 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19 | |
| | | |
| | TACACTATGA GTAAGAAAGG AACAATCGGG GTAACGAGCG ACAATATATT CCCCGTCATC | 60 |
| | AAAAAATTCC TGTACAGCGA CCATGAGATA TTCCTGCGTG AGATCGTCTC CAATGCCGTG | 120 |
| 55 | GATGCTACGC AGAAGCTGAA AACGCTTACA TCCGTCGGCG AATTCAAAGG CGAGACGGGT | 180 |
| 00 | GACCTCCGCG TAACGGTCAG CGTGGATGAA GTGGCACGCA CGATCACGGT CAGCGACCGC | 240 |
| | GGCGTAGGGA TGACCGAAGA GGAGGTGGAG AAGTACATCA ATCAGATTGC TTTCTCCAGT GCGGAAGAGT TTCTTGAAAA GTACAAAGAC GACAAGGCCG CCATTATCGG CCACTTCGGA | 300 |
| | CTCGGATTTT ACTCGGCTTT CATGGTGTCC GAGCGAGTGG ACGTGATCAC GCGCTCTTTC | 360 |
| | CGAGAAGATG CTACGGCGGT GAAATGGAGC TGCGACGGAT CGCCCGAATA CACGCTCGAA | 420 |
| 60 | CCTGCGGACA AGGCTGACCG TGGCACCGAC ATCGTGATGC ACATCGATGA GGAGAATAGC | 480 540 |
| | GAGTTCCTCA AAAAAGAAAA GATAGAGGGG CTCCTCGGCA AATACTGTAA GTTCCTTACC | 600 |
| | GIGCCGATCA TTTTCGGCAA GAAGCAGGAA TGGAAAGACG GCAAGATGCA AGATACGGAC | 660 |
| | GAGGACAATC AGATCAACGA CACACATCCT GCCTGGACCA AAAAGCCTGC CGACCTCAAG | 720 |
| 65 | GACGAAGACT ATAAGGAATT TTACCGTTCG CTCTATCCCA TGTCCGAAGA GCCTCTCTTC | 780 |
| UJ | TGGATCCACC TCAATGTGGA CTATCCGTTC AATCTGACAG GTATCCTCTA TTTCCCGAAG | 840 |
| | ATCAAAACA ACTIGGATCI GCAGCGCAAC AAGATTCAGC TCTACTGCAA TCAGGTTTAC | 900 |
| | GTCACCGATG AAGTACAGGG TATCGTGCCG GACTTCCTCA CCCTCCTGCA CGGGGTCATC | 960 |
| | GATTCGCCGG ATATTCCCCT CAACGTATCG CGCTCCTATC TGCAGAGCGA TGCCAATGTG | 1020 |
| 70 | AAGAAGATCT CGTCTCATAT CACCAAGAAG GTGGCAGACC GTCTGGAAGA AATTTTCAAA AACGACCGCC CCACATTCGA GGAGAAATGG GATAGTCTGA AGCTCTTCGT CGAATACGGT | 1080 |
| | ATGCTGACGG ATGAGAAGTT CTATGAGCGT GCAGCCAAAT TCTTCCTTTT CACCGATATG | 1140 |
| | GACGGACACA AGTACACGTT CGACGAATAC CGAACGCTCG TCGAAGGTGT ACAGACGGAT | 1200 |
| | AAGGACGGAC AGGTAGTGTA TCTCTATGCT ACGGACAAGC ATGGACAGTA CAGCCACGTG | 1260 |
| | AAACGTGCAT CCGACAAAGG CTACAGCGTG ATGCTGGTTGG ATGGTCAGTT GGATCCGCAT | 1320 |
| <i>7</i> 5 | ATCGTGAGCC TGCTGGAGCA AAAGTTGGAG AAGACACACT TTGTCCGTGT CGATAGCGAT | 1380 1440 |
| | - I TOTOGOTOT GUNTAGGAT | 7440 |

| 5 | ACGATCAACA ATCTGATCCG CAAGGAGGAA AGAGCCGAAG TGAAACTGTC CGATACGGAG CGCGCCACTC TCGTGAAGCT CGGAGCCGAA GGTGCACC GGGACGAGAA GAAGCACTTC CAATGTAGCTT TCGAATCGCT CGGAGCCGAA GGTGAAGCCA TCCTTATCAC ACAAGCCGAA CTCCCCGATT CGTACAATCT GGTACTTAAT ACCGATCATC CGTCACAGAC CTCACAGACC TTAGAGCGAA AATCGCCGAG CTGAAAGCGG AAGAATCGGT AGAGCCTTCG CTCACAGAGC TTAGAGCGAA AATCGCCGAG CTGAAAGCGG AAGAATCGC AGGGTACTC GCGAGAAAAAG GGAAGAAACC GGAGGAAATC | 1500 1560 1620 1680 1740 1800 1860 |
|------------|--|--|
| 10 | CCTGTTGCCA CGAAGGAAGC CAAGGAGAAC AACGCCGTCG AACAGGCCAA AACCGAAGGC AGTATCAACG ATCAACTGAC CAAATATGCT CAGGACAACG AGCTGATAGG TCAGCTCATC GACTTGGCTC TGCTCGGAAG CGGATTGCTG ACGGGAGAGG CTTTGGCCGA ATTCATTCGT CGCAGCCAGC GTCTTCTC | 1920 1980 2040 2058 |
| 15 | (2) INFORMATION FOR SEQ ID NO:20 (i) SEQUENCE CHARACTERISTICS: | |
| 20 | (A) LENGTH: 1446 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 25 | (iii) HYPOTHETICAL: NO | |
| 20 | (iv) ANTI-SENSE: NO | |
| 30 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 30 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11446</pre> | |
| 35 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20 | |
| 40 | GACATTCGAC AGAAAAGACC GTGCTTCAAC GCCAATTTGT ACTTTACCG TGCAATGGAA AAACTGATCG ATATTTTGGT CGTAGACGAT GATGTGCAG TCTGTGCCGC ACTGCGTCTG GTGCTCAAGC GAGCGGGCTA TAATCCCGTT ATAGCCAACA GTCCCGACGA AGCTTTGTCC ATAATGCGGA ATCCTGATGG CGCCTGTAAG CCGGCTGTGA TTCTGATGGA TATGAATTTC TCCCTTTCGA CCTCCGGCAG GGAAGGATTG GAACTACTGG AGAAGTACCA GATATTCACT TCCTGCCCTG TCATACTGAT GACGGCTTGG GCTTCGATTC CACTGGCAGT GGAGGGAATG AGGCTTGGAG CTTTCGACTT CATAGGCAAG CCATTGGGACA ACGATCGGCT CCTTCGTACC | 60 120 180 240 300 360 420 |
| 45 | ATAGATACGG CCTTGCATCT GGCTGCTCCC TCAGCTGTGG CGAATCCATC GGAACAGTCT GACAGAGATA CAGCCCGTCA GCCGAAAGCT ACAGTCCAAG AGAATGACCC CTGTGCCCAT ATCATAGGCC GGAGCGATGC CATCTGTAAG ATCAAGGAAC GGATACGCCG CATAGCTCCC ACCCATGCCT CTGTGCTGAT CACGGGCGAA ACGAGCTTACG GCAAAGAGTT GATAGCCGAA GCTCTGCACC GTGGGAGCAA ACGAGCCTCA GCCCCATTCG TCAAGGTCAA TTTGGGTGGG | 480 540 600 660 720 |
| 50 | ATTCCCGAAA GTTTGTTCGA AAGTGAGCTG TTCGGACATA AGAAAGGAGC TTTTACCAAT GCTTTTTCCG ACAGGAAAGG ACGGTTCGAG CTGGCTGATG GCGGCACCAT CTTTCTGGAC GAAATAGGCG ACACTACCGGT CGGCAACCAA GTAAAACTGC TGCGAGTGCT ACAGGAACAG ACATTCGAGC CGTTGGGCGA GAGCGTCTCC CACCGAGTGG ACATCCGTCT GGTATCGGCT ACGAATGCTT CCTTGGAGCG AATGGTAGCC GAAGGACGTT TCAGAGAGGA CCTCTACTAT | 780 840 900 960 1020 |
| 55 | CGATCAACC TGATACATCT GCATCTGCCT CCGCTGCGTG AGCGTCAGGA GGATATACAG CTGCTGGTGG AAGCCTTCAG TGAAGCCTTT GCCCAATCGA ACGGATTGCC CCATGCCGTT TGGAGTGCGG AAGCTATGCG ACGTATCGT GCCATGCCCC TACCGGGCAA TGTACGCGAA CTGAAAAACG TAGTGGAGC TACGCTATTG CTCTCGGGAT CGAGAGAAAT CAGTGCCCG GATGTGGCTG ACTTCGGTTC GCAGGAGAAC ACTCCGACGA ACGGCTTTG | 1080 1140 1200 1260 1320 |
| 60 | ACCGACATGG AGGAAGCTGC TATCCGAGAG ACGCTGACTA AATACAACGG CAACGTTAGT CGTGCTGCAC GAGCCTTGGG ATTGAGCCGG GCAGCTCTTT ACCGGCGAAT GGAGAAATAC GGACTG | 1380 1440 1446 |
| 0.5 | (2) INFORMATION FOR SEQ ID NO:21 | |
| 65 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 789 base pairs | |
| 70 | (B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 7 5 | (iii) HYPOTHETICAL: NO | |

```
(iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 5
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...789
10
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21
      AGATCTCTGC AATCATTTCA AAATAAAAAA CACTCAAGTA TGCTTAAGAT AAAGAACCTC
                                                                                     60
      CACGCCACAG TACAGGGCAA AGAGATATTG AAAGGAATCA ATCTGGAGAT CAATGCCGGA
                                                                                    120
      GAGATTCATG CTATCATGGG GCCGAACGGA TCGGGGAAAA GTACGCTCTC TTCCGTTTTG
                                                                                    180
15
      GTGGGACATC CCTCCTTTGA AGTCACGGAA GGAGAGGTGA CATTCAATGG AATCGACCTG
                                                                                    240
      CTCGAACTCG AACCGGAAGA ACGTGCACAC CTCGGACTCT TTCTCAGTTT CCAATATCCG
                                                                                    300
      GTCGAGATCC CGGGCGTCAG CATGGTGAAT TTCATGAGGG CAGCTGTCAA TGAACATAGG
                                                                                    360
      AAAGCGATCG GAGCAGAACC CGTATCGGCA AGCGACTTCC TCAAGATGAT GCGAGAGAAG
                                                                                    420
      CGTGCCATTG TGGAGCTGGA CAACAAATTG GCCAGCCGTT CTGTGAACGA AGGCTTCTCC
                                                                                    480
20
      GGTGGAGAAA AAAAGAGGAA CGAAATCTTC CAAATGGCTA TGCTCGAACC CAAGCTGGCT
                                                                                    540
      ATTTTGGACG AAACCGATAG CGGGCTCGAT ATCGACGCTC TCCGCATCGT AGCAGGCGGG
                                                                                    600
      GTAAACCGAC TCCGCTCTCC GGAGAATGCT GCTATTGTGA TCACACACTA TCAGCGTTTG
                                                                                    660
      CTCGAGTACA TCAAGCCGGA CTTCGTACAC GTCCTTTACA AGGGGCGCAT CGTCAAGTCG
                                                                                    720
      GGAGGAGCCG AGCTGGCTCT CACGCTCGAA GAAAAAGGCT ACGACTGGAT CAAGGAAGAG
                                                                                    780
25
      ATAGGAGAA
                                                                                    789
       (2) INFORMATION FOR SEQ ID NO: 22
30
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1386 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
35
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
40
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1386
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22
50
      AGCATGGCTA AGGAGAAAAC GATCTACGTC TGCCGTTCGT GCGGAACCAA ATACGCCAAA
                                                                                     60
      TGGCAAGGCA ACTGCAATGC CTGTGGAGAG TGGAACTGCA TTGATGAGGA GAAGGTGCCG
                                                                                    120
      GCACCGGCAT CGGGCAAGCA TGCAGCCAAG AGTTTTATGC CTCGGGAGCA GGACAACCGG
                                                                                    180
      CCAAGACTCT TACAGGATGT GGAGTCCGGC GATGAAGAGC GTATTCGCCT CGGCGATGAA
                                                                                    240
55
      GAGTTCGACC GCGTACTGGG TGGAGGAATT GTCAAAGGAG CATTTGTCCT GCTTGGCGGC
                                                                                    300
      GAGCCGGGAA TCGGTAAGTC CACGCTTATC CTCCAGACGG TGCTGCGTCT GCCGCAGTTG
                                                                                    360
      CGCACGCTCT ATGTGTCGGG CGAAGAAAGT GCCCGACAAC TGAAGATGCG CGCCGAACGA
                                                                                    420
      CTGGGGCAAG CCATGAATGG GTGCTACGTA TACTGCGAAA CGAATATAGA GAGGATACTC
                                                                                    480
      TCCCGTGCAG AAGAACTCAC ACCCGATCTC CTCGTGATAG ACTCTATACA GACGGTCTAT
                                                                                    540
60
      ACCGAGGAAA TGGAAAGCTC GGCCGGCAGC GTGGGGCAGA TCCGCGAATG TGCCGCCTTA CTGCTCAAAT ACTGCAAGAC TACGGGTATC CCCGTCATCG TCATCGGACA CATCACCAAA
                                                                                    600
                                                                                    660
      GAAGGTAGCA TAGCCGGACC GAAGGTGCTG GAGCATATAG TGGATACGGT GCTTCTCTTC
GACGGGGATA AGCATCATCT CTACCGGATA CTCCGAGGAC AGAAGAACCG CTATGGCAGT
                                                                                    720
                                                                                    780
      ACTTCCGAGC TGGGGATATA CGAGATGCGG CAGGACGGTC TGCGTGGCGT GGAGAATCCG AGCGAACATC TCATCACACG CAATAGGGAA GACCTCAGTG GCATAGCCAT AGCCGTAGCG
                                                                                    840
65
                                                                                    900
      ATGGAGGGCA TTCGCCCGAT ACTCATCGAA GCGCAGGCTT TGGTCAGCTC GGCCATTTAT GCCAATCCGC AGCGTTCGGC CACGGGCTTC GATATTCGGC GGATGAACAT GCTCTTAGCC
                                                                                    960
                                                                                   1020
      GTACTGGAGA AACGTGCCGG CTTCAAGCTC ATACAGAAGG ATGTGTTTCT GAACATTGCC
                                                                                   1080
      GGAGGTATCA AAATAGCCGA TCCGGCTACG GATCTGGCCG TTATCTCGGC AGTGCTGGCG
                                                                                   1140
70
      TCGAGTCTGG ACATCGTTAT CCCGCCGGCC GTATGCATGA CGGGCGAGGT CGGACTCTCC
                                                                                   1200
      GGAGAGATAC GTCCCGTGAG CCGCATCGAG CAGCGCATAA CGGAAGCGCG TCGCATAGGG
                                                                                   1260
      TTCAAAGAGA TATTGGTACC GGCCGATAAT TTCCGGCAGG AGGATGCCGG CCGCTTCGGT
                                                                                   1320
      ATTCGGCTCG TGCCGGTCAG AAAGGTGGAG GAAGCCTTCC GCCATCTGTT CTCGAAAGGA
                                                                                   1380
      AGAGAA
                                                                                   1386
75
```

PCT/AU98/01023

```
(2) INFORMATION FOR SEQ ID NO:23
             (i) SEQUENCE CHARACTERISTICS:
  5
                   (A) LENGTH: 1119 base pairs
(B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                   (D) TOPOLOGY: circular
 10
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
            (iv) AHTI-SEHSE: NO
 15
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
20
                   (A) HAME/KEY: misc_feature
                   (B) LOCATION 1...1119
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23
25
       GGGTCTTGTG GAAGTAGCCC AGCAATTGGC CGATCAGGGT GTTCGCGTTG TGATCGCCGG
       ATTGGACATG GACTTTCGAC GTCAGCCTTT CGGACCTATG CCGGGCTTGT GTGCCATAGC
                                                                                     120
       CGACTCCGTG ACCAAAGTTC ATGCCGTGTG TGTGGAATGC GGCCGATTGG CCAGCTATTC
       TTTCCGTCGT GTCCAAGGCG ATCAGCAAGT GATGCTGGGC GAACTGAACG AATACAGTCC
                                                                                     240
       CCTCTGCAGA ACCTGCTACA GGAAATGCAG TTCTCCCCCA CAAACAGAAG AAATCCATTC
30
       GACAATATGA ATAGCAGACA TCTGACAATC ACAATCATTG CCGGCCTCTC CCTCTTTGTA
                                                                                     360
       CTGACATTGG GCGGCTGCTC CGTAGCCCAA CAAGATACGC AGTGGACTCT CGGCGGAAAG
CTCTTTACTT CGGCGTGGAT ACAACGTTCG GCCGAATATC AAGCGCTTTG CATTCAGGCA
                                                                                     420
                                                                                     480
       TACAACATCG CTACGGAAAG AGTGGACGCT CTACCGGCAG AACGTAAACA AGGAGATAGG
                                                                                     540
       CCTTATGCCA TCGTAACGGA CATAGACGAA ACCATTTTGG ACAATACGCC TAACTCCGTG
                                                                                     600
35
       TATCAGGCTC TCAGGGGCAA GGATTATGAT GAAGAGACTT GGGGGAAATG GTGTGCACAG
                                                                                     660
       GCCGATGCCG ACACACTGGC AGGAGCTTTG TCTTTCTTCC TCCATGCAGC GAACAAGGGG
                                                                                     720
       ATCGAGGTCT TTTACGTCAC CAACCGCAGA GACAATCTGC GCGAAGCAAC TCTTCAGAAC
                                                                                     780
       CTTCAGCGTT ACGGATTCCC CTTTGCCGAT GAAGAACATT TGCTTACGAC CCATGGGCCA
TCCGACAAAG AACCCCGTCG GCTCAAAATA CAAGAACAGT ATGAAATAGT ATTGCTCATA
                                                                                     840
                                                                                     900
40
       GGAGACAACT TGGGCGACTT CCACCACTTC TTCAATACGA AAGAAGAGTC CGGACGCAAA
                                                                                     960
       CAGGCTCTGG GCCTGACAGC CGGGGAGTTT GGCCGGCACT TCATCATGCT GCCCAATCCC
                                                                                    1020
       AACTACGGAT CTTGGGAACC GGCATGGTAC GGCGGGAAGT ATCCGCCACT GCCCGAAAGA
                                                                                    1080
       GACAAAGCAC TTAAACAACT GCACTCACAG AACAGCAGA
                                                                                    1119
45
       (2) INFORMATION FOR SEQ ID NO:24
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1278 base pairs
50
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
55
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SEUSE: NO
60
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (iz) FEATURE:
                  (A) NAME/KEY: misc feature
65
                  (B) LOCATION 1...1278
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24
      CGTTGGGGTT TTTCTAATTT TGTGGCCATG AGCACCAATA TAGATGTACA ACAGATCAAA
      CAGCGTTTCG GCATCATCGG TAGCAGTCCG CTGATGGAAC ATGCCATACG AGTGGCAGCA
CAGGTGGCTC CTACCGACAT GTCCGTCCTC GTGACGGGGG AGAGCGGTTC GGGGAAAGAG
70
                                                                                    120
                                                                                    180
      TTCTTCCCAC AGATAATCCA CTACTACAGC GCCCGGAAAC ATCATAGCTA CATTGCAGTC
                                                                                    240
      AATTGCGGAG CCATCCCGA AGGAACCATC GATTCCGAGC TGTTCGGACA CCGCAAAGGT
                                                                                    300
      TCCTTTACCG GAGCCGTATC GGATCGCAAG GGGTACTTCG AAGAAGCATC CGGCGGCACG
                                                                                    360
75
      ATCTTTCTGG ACGAAGTGGG CGAACTGCCT TTGCCCACGC AGGCGAGGCT GCTGAGGGTG
                                                                                    420
```



| 5 10 15 | CTGGAGACGG GCGAGTTCAT CCCCGTAGGA GCCAGCCAGT CGCAGAAGAC GGATGTCCGT ATCGTAGCGG CGACGAATGT GAACCTCAAG GAGCGGTAG CGAACGGGAA GTTCCGGGAA GACCTCTTCT TCCGGCTCAA TACGGTACCG ATCGAGGTC CTGCCTGCG TATGCGAC CCTCCGCTGC GCCTATCGGA CGCAGCACA ACCATATTAA TGCGTTACCG CGAACGCCAA TATAACCGAC AGCGTGAGCA TCCTGGAGGA GGAGCGGACA GCGCCGGAGA GTATCGGACG GTATCGGACA AGCGACCACA AGCGACCAA AACGACCAA AGCGACAACAAC AAATCCCCCA TTACGAGCGC AACGAACACAC CGAACAGCC TCATGGCTT TAGGGTCGGA CGTCTGGGGCACA AGCGACACAC CGAACAGCC TCATGGCTT TAGGGTCGGA CGTCTGGGGC AACGACCAC AGATCCACCCC TCATGGCCT TCAGCGCACA GCCCCACACA GCCCCCACACACACACACA | 480 540 660 720 780 960 1020 1080 1140 1260 1278 |
|---------------|--|--|
| | (2) INFORMATION FOR SEQ ID NO:25 | |
| 20 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1959 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 25 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | | |
| 20 | (iii) HYPOTHETICAL: NO | |
| 30 | (iv) ANTI-SENSE: HO | |
| | <pre>(vi) ORIGINAL SOURCE: (A) ORGANISM: FORYPHYROMOMAS GINGIVALIS</pre> | |
| 35 | (ix) FEATURE: | |
| 00 | (A) NAME/KEY: misc_feature (B) LOCATION 11959 | |
| 40 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25 | |
| | AAAAATCTTC CGAGGTACGC TCCGGCATTT GGCAGAAATC GTAAAAGGGA ACAACCTGAC CCTGACAACG ATGATAGTCG TGGGAGAAGC CATAGACAAT CGGGAAGGGC TGTCACGGTT GTATGCCGAC GAATTCAAAC ACCTCTTCCG CACATGATCC TACTCTTCGG CGGTACTACG GAAGGCCGTG CCGCAGCTCG CGTGCTGGAT GAAGCGGGAA GTCCGTTTTT CTACTCCACC | 60 120 180 240 |
| 45 | AAAGGCAATC TGCAAGAGAT CCAGAGTAGC CACGGCCATC GTCTGACAGG AGCCATGACG GTTGCCGACA TGGTTTCGTT TTGTCGGAAA GAAGAGATCC GACTGATCGT GGACGCCGCT CATCCTTTCG CCGAAGAATT GCACGCTTCA GTGGCAGAAG CCTCTGAACA AACAGGTATC CCCGTAGTAA GATACGAGAG ACAATACCCT CCACGCGAAG AAGGTATCGT CTGGTGTGCA | 300 360 420 480 |
| 50 | AACTACGATA CGGCTGCCGA GCGGATGCTT GGCGATGGCG TGCAGCGTCT GCTGATGCTC ACAGGAGTGA ATACGATCCC CAAGCTGGCT GCTTTCTGGA AAGAGCGCAC CACCTTTTGC | 540 600 |
| | CGCATATTGA AGCGAGACGA ATCGGTTGCT TTGGCAGAGA AGAACGGCTT TCCTGCGGAG CGCATCGTTT TCTTCGAACC GCATGCGGAC GAGGAGCTGA TGCAAGCCGT TCGCCCCGAT | 660 720 |
| | GCCATTATCA CAAAAGAAAG CGGAGAGAGGC GGTTACTTCC GAGAAAAGAT AGAAGCTGCC | 780 |
| 55 | CGACGGATGG GCATCCGTAT ATATGCCGTC GTACGTCCCC CTTTGCCTCC TTCATTCATT CCCGTAGGCG GGCCTGTCGG TTTGAGACGG GCGGTAGAAC GCCTCGTGCC GGGATTCTTT | 840 900 |
| | TCACTCCGAA GCGGATTCAC TACCGGCACC ACAGCTACCG CTGCAGTAGT AGCAGCCATG TACCGATTGA TGGGGCTTGG CTCTCTCGCC GAAGCTCCCG TAGAATTGCC TTCGGGCGAA | 960 |
| | ATAGTCAGTC TGCCCATAGC GGAAATTCGA GAGGAAGAAG ATGCTGTCGT ATCCGCAGTC | 1020 1080 |
| 60 | CTGAAAGATG CAGGTGATGA TCCGGATGTG ACCAATGGCA TGGCGGTATG CGCTACGATC AGGCTCAATC CCGAACATGA GGAAGTCCGC TTCCTGCAGG GTGAAGGGGT GGGGGTAGTG | 1140 1200 |
| | ACGCTCCCCG GCCTCGGTCT GGAGGTCGGA GGTCCGGCTA TCAACCTCGT ACCTCGACGA | 1260 |
| | ATGATGACAG CAGAGGTACG CCGACTCTAT GCGCAGGGAG GTGTGGATAT TACGATTAGC GTACCCGAAG GCCGAGAGGC TGCTACCCAG ACATTCAATC CCCGACTCGG CATACGGGAC | 1320 1380 |
| 65 | GGCATCTCTA TTATCGGAAC ATCGGGAGTC GTGAAACCTT TTTCGGCCGA AGCGTTCGTT GGTGCCATCC GTAAGCAAGT GGGTATTGCC ACCGCCTTGG GAGCCAATCA TATCGTCCTC | 1440 1500 |
| | AATTCGGGAG CCAAGAGTGA GCGTTATGTA AAAGGAGCCT ATCCGGCACT CATTCCACAG | 1560 |
| | GCCTTTGTGC AGTATGGCAA TTTCGTCGGC GAATCACTCA GTTGTGTAGC TTCCTTCCCT TCTGTCCGTT CGGTAACGGT AGGAATCATG CTCGGCAAAG CAGTGAAACT CGCCGAAGGC | 1620 1680 |
| 70 | TATCTGGATA CGCACAGTAA AAAGGTAGTG ATGAATCGGG ATTTCCTGCA CGAACTGGCT | 1740 |
| 70 | CGTCAGGCAG GTTGTTCGGA AGACATCCAT GCCATAATAG ACAGCCTGAA TTTGGCTCGT GAGCTATGGA CTATGCCGAG TGCGGAGGAC AGCGATCGAC TGCTACGAAA GATTGCCGAA | 1800 1860 |
| | CGATCTTGGG AAACTTGCCG CCCATCGGTA CCATCGGCCG AATTAGAACT CCTGCTGATC GATGAGTCCG GAGCGATTCG TTTTCGTATC GGTGGAGAA | 1920 1959 |
| | | |

75

```
(2) INFORMATION FOR SEQ ID NO:26
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1353 base pairs
  5
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
            (ii) NOLECULE TYPE: DNA (genomic)
10
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
15
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
20
                  (B) LOCATION 1...1353
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26
       CAACAAAGAC CACATAGAAT TATTATGTTG AGGACTTTCC GAATCGGTGG TATTCACCCC
                                                                                   60
25
       CCCGAAAACA AGTTGTCGGC AGGCAAGCCC GTAGAGGTGT TGCCTATCCC CTCACAGGTA
                                                                                  120
       GTCATCCCTC TTGGTCAGCA CATCGGTGCA CCGGCAACTG CCACGGTCAA GAAAGGGGAT
                                                                                  180
       GAAGTTAAGG TCGGGACTAT CATTGCTCAG GCCGGAGGAT TCGTATCAGC TAATATCCAC
                                                                                  240
       TCATCTGTGT CGGGTAAGGT GCTGAAGATC GATAACGTAT ACGACTCAAG CGGCTATCCC
                                                                                  300
       AAGCCCGCAG TCTTCATTAG CGTAGAAGGT GACGAATGGG AAGAGGGCAT CGATCGCTCA
                                                                                  360
30
       CCAGCCATCG TCAAAGAATG CAATCTGGAT GCAAAAGAAA TCGTAGCCAA AATTTCTGCA
                                                                                  420
       GCCGGTATTG TGGGTCTTGG CGGTGCTACC TTCCCTACCC ATGTGAAGCT GTCCCCTCCT
                                                                                  480
       CCGGGCAACA AAGCTGAGAT CCTGATCATC AACGCCGTAG AGTGCGAGCC TTATCTGACG
                                                                                  540
       AGCGACCATG TCCTTATGCT GGAGCACGGC GAAGAGATCA TGATCGGCGT GAGTATCCTG
                                                                                  600
       ATGAAAGCCA TTCAGGTAAA CAAGGCCGTC ATCGGAGTTG AGAATAATAA GAAAGATGCT
                                                                                  660
35
       ATTGCTCACC TCACCAAACT GGCCACTGCA TATCCGGGCA TAGAGGTAAT GCCGTTGAAG
                                                                                  720
       GTGCAATATC CTCAAGGCGG TGAGAAGCAG CTGATCGATG CAGTGATCCG CAAGCAGGTA
                                                                                  780
       AAAAGCGGTG CCTTGCCTAT CAGCACAGGT GCCGTAGTAC AAAACGTGGG TACGGTATTC
                                                                                  840
       GCCGTGTACG AAGCAGTACA GAAGAACAAG CCTCTGGTCG AGCGCATCGT GACGGTTACA
GGAAAAAAAC TGTCTCGTCC GTCTAACCTC CTCGTTCGTA TAGGTACTCC TATTGCGGCT
                                                                                  900
                                                                                  960
40
       TTGATCGAAG CAGCAGGTGG CTTGCCGGAG AATACGGGCA AGATCATCGG CGGAGGTCCG
                                                                                 1020
       ATGATGGGAC GCGCTCTGCT GTCACCGGAT GTGCCTGTGA CCAAAGGCAG CTCCGGAGTA
                                                                                 1080
       TTGATTCTCG ATAGAGAAGA GGCAGTTCGC AAGCCTATGC GCGACTGTAT CCGATGCGCC AAGTGCGTCG GAGTGTGTCC GATGGGACTC AATCCGGCTT TCCTTATGCG CGACACCTTA
                                                                                 1140
                                                                                 1200
       TATAAGAGCT GGGAAACAGC GGAAAAAGGC AACGTGGTTG ACTGTATCGA ATGCGGTTCG
                                                                                 1260
45
       TGCAGCTTCA CCTGTCCGGC CAACCGTCCT CTGCTGGATT ATATCCGCCA AGCCAAGAAG
                                                                                 1320
       ACTGTGATGG GTATCCAAAG AGCACGTAAG CAA
                                                                                 1353
       (2) INFORMATION FOR SEQ ID NO: 27
50
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1467 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
55
                 (D) TOPOLOGY: circular
           (ii) HOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
60
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORTPHYROMONAS GINGIVALIS
65
           (ix) FEATURE:
                 (A) HAME/KEY: misc_feature
                 (B) LOCATION 1...1467
70
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27
      ATGAACTCTC AAAAGAAAGA GGCTTTTAAT ATGAAAAGAA TACAACTAAC TCTTATCGCT
                                                                                   60
      CTCTTCGCCG CTGTTGCCGG TTTGGTCGCT CAAAATGCTT ACGAGGGAGT AATTTCATAT
                                                                                  120
      AAAATTTCGT TGGACAAAAC CGGAAACAAG GTTGTACTGA ATGGTGCGGC AGATATGAGT
                                                                                  180
75
      AATTTAAAGC TCAAGAGCAC TCAGATGATC ATTGTTACGC CTATTCTTCG TTCAGAAGAT
                                                                                  240
```

PCT/AU98/01023

WO 99/29870

```
GGTACCAGCC GGGTGGAATT TCCTTCGGTA GTCATTACAG GCCGCAATAG AACAAAAGCT
                                                                              300
      CTCAAGCGTG AAATCGCATT TAGTTCGGCT TTGCCCCAAG CAAAACATGC AGCTCAATAC
                                                                              360
      ATTCGCCGTC ATAATGGGAA GAGCGAGCAG TTTGCTTTTA CAGGAGAACA TGCTTATGCA
                                                                              420
      TCATGGATGA TGGATGCCAA GTTTGTGGTT CGTGAGGAGG TACGAGGTTG TGCTAAATGC
                                                                              480
 5
      CCTGTAGGTC TCTCGAGTAA TATTGTTCCT TTTGATCCAC TCTTCAATCC GGCAGAGGCT
                                                                              540
      CCTTATTTGT TGGCACACAT TACTCCGGCA GAAGAAGTGG AAAAACAGCG AGAGTCCAGC
                                                                              600
      TTCGATGCTT ATATCAACTT CAAAGTCAAT AAGGCAGATG TCCTTCCTGA GTATCGCAAC
                                                                              660
      AATAAGGCGG AGTTAGAGAA AATCAAAGAA TTTGTAAGCA CCGTTAAGGC TAATCCAAAC
      TATTCGGTCA ATAAAATGAT CATCGAAGGG TTTGCTTCTC CCGAGGCTTC AATAGCCCAC
                                                                              780
10
      AATAAGGCTT TGTCGGAGCG CCGTGCTAAA AGACTCGCGG AAGAATTGGT GCGTAAGTAT
                                                                              840
      GGCAAAACAT TGCCGAATAT AACCACTGAA TTCGGCGGTG AAGATTGGAA GGGGCTGAAA
                                                                              900
      CTGGCTATCG AAAAGAGTGA TATAGCCGAT CGTGACCGCG TATTGGAGAT AATCAACTCC
                                                                              960
      GATAAATATG CCGATGATGA TGCACGTGAA CAGGCTCTGA AGCAACTTTC GTCTTATCGT
                                                                             1020
      TATATCTTGG ATCAGATCTA TCCGAATTTG CGTCGCAATA CGATAACCAT GGGGTATATC
                                                                            1080
15
      GTTCGTGATT ATACCCTCGA AGAAGCTCGT GAAATCATTA AGACTGCTCC GAAAGAACTT
                                                                             1140
      AGTGAGGCCG AAATGTACCG TGTGGCAATG TCTTATCCTG AGGGGCACCA AGAGCGTTTG
                                                                            1200
      TTTGCTCTGA ATACGACCCT TAAGTATTTC CCTGAAAGTG TAACGGGCCG AATCAATTTG
                                                                             1260
      GCTGTAGCCG CTTTTAATGG TGGAGACGTT CAACAGGCAA TTGCTCTGTT GAGTCCGATT
                                                                            1320
      CAGACAGAAA AGGGTGTAAG CAATATCCTT GGAGCTGCTT ATGCTCGTAC GGGAGATTTT
                                                                            1380
20
      GCTCGTGCCG AAACCTTCTT CCGTAAGGCC GTTGCAGAAG GAGATGCAAA TGCGCAGCGC
                                                                            1440
      AACCTCGATA TGCTGCTTGG CAAAAAG
                                                                            1467
      (2) INFORMATION FOR SEQ ID NO:28
25
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1152 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
30
                 (D) TOPOLOGY: circular
          (ii) HOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
35
          (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
40
          (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1152
45
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28
      GACATGGCAG AAAAAAGAGA CTATTACGAA GTCCTCGGTG TATCGAAGAA TGCCACCGAC
      GATGAACTGA AAAAAGCATA TCGCAAGAAG GCTATCCAAT ACCATCCTGA TAAGAACCCC
                                                                             120
      GGTGACAAGG AGGCCGAAGA GCACTTCAAA GAGGTAGCTG AAGCCTACGA CGTATTGAGC
                                                                             180
50
      GATCCGCAGA AGCGCAGTCA ATATGACCAG TTCGGCCATG CCGGATTGGG CGGAGCTGCC
                                                                             240
      GGTGGAGGTT TCAGCGGAGG CGGTATGTCC ATGGAGGATA TTTTCAGTCG CTTCGGTGAT
                                                                              300
      CTATTCGGTG GGTTCGGCGG TTTCGGCGGA TTCTCCGATA TGGGCGGTGG CAGTCGCAGA
                                                                              360
      CGTGTTCGCA GAGGGTCTGA CCTGCGAGTA CGAGTGAAGC TTTCTTTGGC CGATATAAGT
                                                                              420
      AAAGGTGTGG AGAAGAAAGT GAAGGTAAAA AAGCAGGTAG TGTGCAGCAA ATGTCGTGGC
                                                                              480
55
      GATGCCACGG AAGAAGCCAA TGGCAAGACT ACCTGCCAGA CCTGCCATGG AACCGGCGTG
                                                                             540
      GTTACACGTG TGAGCAACAC TTTCCTTGGG GCCATGCAGA CCCAGAGCAC TTGTCCCACT
                                                                             600
      TGCCACGGAG AGGGTGAGAT CATCACGAAG CCATGCTCCA AGTGTAAGGG CGAAGGTGTG
                                                                             660
      GAGATCGGCG AAGAGGTGAT CTCATTCCAC ATCCCTGCCG GTGTAGCCGA AGGAATGCAA
                                                                             720
      ATGTCCGTGA ACGGCAAGGG AAATGCCGCG CCCCGAGGAG GCGTGAATGG CGACTTGATA
                                                                             780
60
      GTCGTGATCG CCGAGGAACC GGATCCGAAT CTGATCCGCA ATGGCAACGA TCTGATATAC
                                                                             840
      AATCTGCTTA TATCCGTTCC GTTGGCTATA AAAGGAGGTA GTGTGGAAGT GCCGACGATA
                                                                             900
      GACGGACGAG CCAAGATCCG CATCGAGGCG GGGACACAAC CCGGCAAGAT GCTGCGTTTG
                                                                             960
      CGCAATAAGG GGTTGCCCAG CGTAAACGGC TATGGCATGG GAGACCAACT GGTGAATGTC
                                                                            1020
      AATGTCTATA TCCCCGAATC GATCGATGCC AAAGATGAGC AGGCTATCGC AGCGATGGAA
                                                                            1080
65
      AACTCGGACA GCTTCAAACC TACCGATGCT GCTCGTAAGG ATATAGACAA GAAATACAGA
                                                                            1140
      GAGATGCTGG AT
                                                                            1152
      (2) INFORMATION FOR SEQ ID NO:29
70
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 927 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
75
                (D) TOPOLOGY: circular
```

```
(ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
  5
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 10
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...927
15
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29
       CGATTCGATT CAACAACTAA TGTCTCACAA ATTAATTTAA GAACAGAGAT GAAAAAACTG
                                                                                     60
       ATTTTAGCGA CTTTGGGACT TATGGCCATT GCCATGCTCT CATGTTCAAG CAACAACAAG
                                                                                    120
       GATTTGGAGA ACAAAGGGGA GGCTACTCTT TTGGTAACGT TTGGTAGCTC CTATAAAGCT
                                                                                    180
20
       CCACGCGAAA CCTATGCGAA GATTGAGAG ACTTTTGCCG CAGCTTATCC CGATCAAAGG
ATAAGCTGGA CATACACGTC TTCTATTATC CGAAAGAAAA TGGCTCAGCA GGGTATTTAT
ATCGATGCTC CGGATGAGGC TTTGGAGAAA TTGGCTCGTC TGGGTTATAA GAAGATCAAT
GTACAGAGTC TTCATGTGAT TCCCGGCCGA GAATATGATC AGATGATCGA CTTTGTCAAT
                                                                                    240
                                                                                    300
                                                                                    360
                                                                                    420
       AAGTTTAAGG CAGCACATAG TGATATTACT GTGAAGGTAG GGGCTCCGCT TTTCGATACC
                                                                                    480
25
       GATGAAGATA TGCGCGAGGT GGCAGAGATC TTGCACAAGC GTTTTCAGCA AACGATAGAG
                                                                                    540
       AAAGGTGAAG CTATTGTATT CATGGGACAC GGCACCGAGC ATGCTGCCAA TGACAGGTAT
                                                                                   600
       GCCCGTATCA ATAAGATCAT GAAGAACTAT AGCAAGTTCA TGATCGTCGG AACCGTCGAG
                                                                                    660
       TCCGATCCCT CTATCAATGA TGTTATTGCC GAACTGAAAG AAACCGGTGC CACGGCCGTA
                                                                                   720
       ACAATGATGC CGCTGATGAG TGTGGCAGGC GACCATGCTA CGAATGATAT GGCCGGAGAT
                                                                                   780
30
       GAGGACGATA GCTGGAAGAC GTTGCTGACC ANTGCCGGCT ACACAGTTTC TATAGACAAG
                                                                                   840
       CTGGACAATG GCAATTTCTC AGCTCTTGGA GATATAGAAG AGATCCGGAA TATCTGGCTC
                                                                                    900
       AAGCATATGA AAGCCACCTC TGCTCGC
                                                                                   927
35
       (2) INFORMATION FOR SEQ ID NO:30
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1473 base pairs
                  (B) TYPE: nucleic acid
40
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
45
          (iii) HYPOTHETICAL: NO
           (iv) AHTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
50
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1473
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30
       CGGAGAGATG CGCGGAGGCA ACTGGTACGC CGTAGGAGGA AAGAGCTATC TGGCACAGCA
       AATACGCGAT GCCGGAGGAG AGTATTTCCT GAAAGACGAT CAGCGATCCG GTGGTGTTTC
60
       CCTCGACTTC GAGACGGTCT ACAGCCGATC CGATTCGGCT CGCTACTGGC GAATCCTCAA
                                                                                   180
       CAGCTATCCC GGGGAGTTCG GTTACGAAGC TTTGAAAGCC GAAGACAGCC GATATGCCGA
                                                                                   240
      TTTCAGGGCA TTCAAGGAAA AAGGCGTGAT ATACTGCAAC CTGAGGGAAA CAGCCTTCTA
                                                                                   300
      CGAACTCATG CCCATGCATC CCGATTGGGT GCTGGCCGAT CTTATCGCTA TCTTGCACCC
                                                                                   360
      CGGACTACTT CCCGACCACC AACCGCATTT CTATTATTTG CTCCAATGAC ATCCGTCAGC
                                                                                   420
65
      CACTTACGTA CAATTTCTGT CGCAGGTATC CTGGCTGCGC TGGGAGGGGC TGTACTCATT
                                                                                   480
      CTCTTCGGGG TTAATCTCTT CCTCGGCTCG GTGGCTATTC CGATGAGCGA GATCTTCCGA
                                                                                   540
      CATCTTTTT CAGATCGTCC CGAAGGAGGA GAAGCACTCG TGCACTACAA TATCCTATGG
                                                                                   600
      AAATCCCGCC TGCCCGAAGC CCTCACGGCT GCTTTTGCCG GCGCAGGTTT ATCCGTTAGT
                                                                                   660
      GGCTTGCAGA TGCAGACCGT CTTTCGCAAT CCTTTGGCCG GTCCGTCCGT TCTCGGCATC
                                                                                   720
70
      AGCTCCGGTG CCAGTTTGGG TGTTGCTTTG GTCGTTCTGC TGAGCGGCTC GCTGGGAGGA
                                                                                   780
      GTGGCATTGA GTAGCCTGGG TTATATGGGC GAGGTGGCCA TGAATATAGC CGCTGCCGTA
                                                                                   840
      GGCTCGCTGG CAGTAATGGG GCTGATCGTT TTTGTCAGCA CCAAGGTGCG CAGCCACGTT
                                                                                   900
      ACGCTGCTCA TTATCGGCGT TATGATCGGA TATGTAGCCA CTGCCGTCAT CGGGGTATTC
                                                                                   960
      AAGTTTTTCA GTATCGAAGA AGATATTCGG GCATACGTAA TTTGGGGGTT GGGCAGCTTT
                                                                                  1020
75
      TCCCGTGCCA CGGATTCGCA ACTGAGTTTC TTTGCCATTC TGATGTTGAT CTTTATTCCG
                                                                                  1080
```

PCT/AU98/01023



20 / 490

```
1140
      GCCGGTATGC TCCTTGTCAA GCAGTTGAAT CTCTTATTGC TGGGAGAAAG CTACGCACGT
      AATCTGGGAC TGAATACTCG TCGGGCACGG CTGCTCGTGA TCTCTTCCGC CGGTTTGCTC
                                                                            1200
                                                                            1260
      ATCGCTACCG TCACGGCCTA TTGCGGTCCC ATCGGCTTTT TGGGGATGGC TGTGCCACAC
      TTGGCACGGG TTATCTTTCA CACATCGGAT CATCGGATCC TGATGCCTGC TACCTGTTTG
                                                                            1320
 5
      ATTGGAAGTG CTCTGGCTCT TTTCTGCAAT ATCATTGCTC GTATGCCGGG GTTTGAGGGG
                                                                            1380
      GCTTTGCCCG TCAATTCCGT AACGCTTTG GTGGGAGCAC CTATTATCGT CACCGTTTTG
                                                                            1440
      TTCCGGCGCA GACGCTTCAA GGAAGAAACC GAC
10
      (2) INFORMATION FOR SEQ ID NO:31
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 2289 base pairs
                (B) TYPE: nucleic acid
15
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
20
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
25
                (A) ORGANISH: PORTPHYROHONAS GINGIVALIS
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
                (B) LOCATION 1...2289
30
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31
      CATTTTTTAG TAACGATTAT GCGGACAAAA ACTATCTTTT TTGCGATTAT CTCTTTTATT
      GCTCTATTGT CGTCTTCTCT GTCGGCTCAG AGCAAAGCCG TTTTAACCGG TAGTGTGTCG
                                                                             120
35
      GATGCCGAAA CCGGAGAGCC TCTTGCCGGT GCTCGAATCG AAGTCAAACA CACCAACATA
                                                                             180
      GTAGCCGGTG CCGATGCCGG CGGACATTTC GAGATCAAGA ACCTGCCGGC AGGGCAGCAT
                                                                             240
      ACTATTATAT GTTCGTTGGG GGGGTATGGA CAGAAAGAGG AGGTGGTTGC CATCGAAGCC
                                                                             300
      GGACAGACCA AAACGATCTC TTTTGCATTG CGACTGCGAA CGAACAACTT GGAGGAAGTC
                                                                             360
      GTCGTTACCG GTACCGGTAC ACGTTACCGC TTGGTCGATG CTCCTGTGGC AACGGAAGTC
                                                                             420
40
      CTTACCGCTA AGGACATAGC CTCTTTCTCG GCTCCTACTT CCGAGGCCTT ATTGCAGGGG
                                                                             480
      CTGAGTCCGT CTTTTGACTT CGGCCCCAAT CTGATGGGCT CTTTCATGCA GCTGAACGGC
                                                                             540
      CTTAGCAGTA AGTATATCCT CATCCTTATC GATGGTAAGC GTGTGTACGG CGATGTAGGC
                                                                             600
      GGTCAGGCCG ATTTGAGTCG TATTTCTCCT GATCAGATCG AACGGATCGA ACTGGTGAAA
                                                                             660
      GGTGCTTCGA GTTCGCTCTA CGGATCCGAT GCCATCGCCG GGGTAATCAA TGTGATCACA
                                                                             720
45
      AAAAAGAATA CGAATCGACT GAGTGCATAT ACGTCACATC GCATATCGAA GTACAACGAT
                                                                             780
      CGGCAAACCA ATACTTCGCT CGATATAAAC ATCGGTAAGT TCAGTAGCAA TACCAACTAT
      TTCTTCTACC ATACGGATGG CTGGCAGAAT AGTCCGTTCG AAATAAAAAA GAAAAAAGGA
                                                                             900
      TCCGGCGAAC CGGTCTTGGA GGAAACGTAT AAGAAAACTT TTCGTGCACA GGAAAATCAG
                                                                             960
      GGTGTAAGCC AATCGCTTTC CTATTATGCA ACTAACAATC TTAGCTTCAG CGGAAATGTG
                                                                            1020
50
      CAGTACAATA AACGTCAGAT CTTCACTCCG ACTTTTTCCG AAAAGAAGGC CTATGACATG
                                                                            1080
      GATTATCGTG CTTTGACGGC TTCACTCGGT ACGAACTATC TTTTCCCCAA TGGTCTGCAT
                                                                            1140
      ACGCTTTCTT TCGATGCCGT CTACGATCGC TTCCGTTTCG GATATTTGTA TCATGACAAG
                                                                            1200
      GACAGCAGTG AGAGCCTGAT CAACAACCAA GGTCAGACCG AGCAACCCAC ATTCTTTCCG
                                                                            1260
      GGTCAGCTAC GCAATAAAAA CGATCAGATC CGATACACGG CAGAGGCTCG CGGTGTATTT
                                                                            1320
55
      ACACTGCCTT ATGCGCAGAA ACTGACCGGC GGTTTGGAGT ATTTCCGTGA GGAATTGATC
                                                                            1380
      TCTCCCTATA ATTTGATTAC CGACAAGGCA GATGCTTCCA CGCTCTCTGC TTATGTACAA
                                                                            1440
      GATGAATGGA AACCGCTCGA TTGGTTCAAT ATGACAGCCG GTTTCCGTCT GGTACACCAT
                                                                            1500
      CAGGAGTTCG GTACACGAAT GACGCCTAAG GTATCCATAC TCGCCAAGTA TGGGCCGCTG
                                                                            1560
      AACTTCCGCG CTACGTATGC TAACGGCTAT AAGACTCCCA CGCTGAAAGA GCTTTTTGCA
                                                                            1620
60
      CGGAACGAAC TCACCACTAT GGGTTCGCAC AATCTCTATC TCGGCAATGC GGATCTTAAG
                                                                            1680
      CCACAGATGT CGGATTATTA TGCTTTGGGC TTGGAGTACA ATCAAGGCCC TATCTCGTTC
                                                                            1740
      AGTGCAACGG TTTATGACAA TGAACTTCGC AATCTGATCT CCTTTATGGA TATACCGACC
                                                                            1800
      TCACCCGAGO ACGAAGCTCA GGGAATCAAG AAAACCAAGC AGTATGCCAA CATAGGAAAA
                                                                            1860
      GCTCGCAGCC GCGGCCTTGA TGTCCTATGT GATGCCTCTA TCGGTTGGGG TATCAAGTTA
                                                                            1920
65
      GGAGCCGGAT ACAGCCTCGT GGAAGCTAAG AATCTCCAGA CGGATGAGTG GCTGGAAGGA
                                                                            1980
      GCTGCACGTC ATCGTGCCAA TGTGCACGCC GATTGGGTTC ACTACTGGGG TCAGTATAGA
                                                                            2040
      CTTGGCGTGA GCCTTTTCGG CCGTATTCAG AGCGAGCGTT ACTACAAAGA CGGCAATGCT
                                                                            2100
      CCGGACTATA CCTTGTGGCG ACTCGCCACA TCGCATCGTT TCGCTCATTT CCGCCACATC
                                                                            2160
      ATCCTGGATG GAACGCTCGG TATAGACAAC CTGTTTGACT ACGTGGATGA TCGTCCTATG
                                                                            2220
70
      GGTGTCAATT ATGCTACCGT AACGCCGGGA CGTACTTTCT TTGCTCAAAT AGCGATTCGA
                                                                            2280
      TTCAACAAC
                                                                            2289
```

(2) INFORMATION FOR SEQ ID NO: 32

```
(i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1095 base pairs
                 (B) TYPE: nucleic acid
                     STRANDEDNESS: double
 5
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
10
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
15
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1095
20
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32
      TGCTGCGCAA GGCTTCGGGA CACCTCTCCC CGGAAGAAGT CGTTTCGGAA TGATACCGTT
      TTTCCTTATT TCGTTATTCA TCTGATCAAA CATATTATCA TTATGACGGA CAACAAACAA
                                                                                120
      CGTAATATCG TATTCCCGGC GTTTCTCCTC TTGCTGGGAG TCATCGCAGT GGTGACGATC
                                                                                180
25
      GTTGGTTTTT TCATGCTCAG ACCGCCCGAG GAGATTATCC AAGGACAGAT AGAAGTGACC
                                                                                240
      GAATACCGAG TGTCCAGCAA AGTGCCCGGG CGCATCAAGG AACTTAGGGT ATCCGAGGGA
                                                                                300
      CAGCAGGTGC AGGCCGGCGA TACCCTCGCT GTCATCGAAG CCCCCGACGT AGCGGCTAAG
                                                                                360
      ATGGAGCAGG CAAAGGCTGC CGAAGCAGCT GCACAGGCTC AGAACGCCAA GGCTCTCAAA
GGAGCACGCA GCGAACAGAT ACAGGCAGCC TATGAGATGT GGCAGAAAGC TCAGGCCGGC
                                                                                420
                                                                                480
30
      GTAGCCATAG CGACCAAGAC ACACCAGCGC GTGCAGAACC TCTATGACCA GGGAGTGGTA
                                                                                540
      CCGGCTCAGA AGTTGGACGA AGCCACTGCC CAGCGCGATG CGGCCATCGC TACGCAAAAA
                                                                                600
      GCGGCCGAAG CCCAGTACAA TATGGCTCGC AACGGTGCCG AACGCGAAGA CAAGCTGGCA
                                                                                660
      GCTTCTGCCC TCGTCGATAG AGCGAGAGGA GCCGTCGCCG AGGTGGAGTC GTACATCAAC
                                                                                720
      GAAACCTACC TCATCGCCCC ACGGGCAGGC GAAGTGTCGG AGATATTCCC CAAAGCCGGC
                                                                                780
35
      GAACTCGTAG GTACCGGCGC ACCTATCATG AATATCGCCG AGATGGGCGA TATGTGGGCC
                                                                                840
      AGCTTTGCCG TTCGTGAGGA TTTCCTCAGC AGCATGACCA TGGGAGCCGT TCTGGAGACT
                                                                                900
      GTGGTGCCGG CTCTGAATGA AGAAAAAGTA CGCTTCAAGA TCACATTCAT CAAGAACATG
                                                                                960
      GGTACCTATG CTGCCTGGAA AGCGACCAAG ACAACAGGGC AGTACGACCT GAAGACCTTC
                                                                               1020
      GAGGTAAAGG CCACCCTTGC GGATAAAGAC AAGGCACAAA AGCTACGCCC GGGTATGTCC
                                                                               1080
40
      GTGATCATAC GCAAG
                                                                               1095
      (2) INFORMATION FOR SEQ ID NO:33
45
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 960 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
50
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
55
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
60
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...9\overline{6}0
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33
65
      CCGCAATCCT CTCCTGATCG AAGAAGCTTC CAAAACGTCA TGAATAAATA TCATTCTCAA
      AGCGTTTTAG AGGTCGGCAA AATTGGGATT GTGATTATCT TTGCGCCCAT AGTACGGAAT
                                                                                120
      GTACATCAAC AACCCCTTT TTTAAGCCAT AAATCAATTA TGCGTATTGT CAGTAATTTT
                                                                                180
      TTGTTCGTCT CTTTTTCGGT TTTGCTTTTT GCATCATGCC GTTCCCAGCG AGAAAAGGTC
                                                                                240
70
      GTTTACCTGC AAGATATCCA AACTTTTAAT CGGGAGATTA TCGCTAAACC ATATGACGTA
                                                                                300
      AAAATTGAGA AGGACGATGT GCTGAACATC CTTGTCAGCA GTAGAGACCC GGAGCTTTCA
                                                                                360
      ACGCCCTACA ACCAAGTGTT GACCACTCGT GCACTGGCCC GCAACGGCTA TGGAACGAAC
                                                                                420
      TCGAACGAAG GCTTCCTGGT CGATTCGAAA GGGTACATCA ATTATCCTAT TTTAGGCCAG
                                                                                480
      ATCTATGTAG AGGGCCTTAC TCGTACCGAA CTGGAGAAGG AGATACAGAA GAGGATTATT
                                                                                540
75
      TCCAGTGGAT TTATCAAGGA TCCTACGGTA ACGGTGCAGC TTCAAAATTT CAAGGTGTCG
                                                                                600
```

PCT/AU98/01023

WO 99/29870

| 5 | GTTTTGGGAG AGGTGAATCA TCCGGGTTCG ATGTCGGTAA AAGGAGAGCG AATAACTCTT TTGGAAGCGA TCGGAATGGC CGGAGACCTG ACAATCTATG GTCGCCGCGA TCGGGTTTTT GTGATTAGAG AAACCGATGG GCATCGCGAG GTTTTCCAGA CGGATCTCAG AAAGGCCGAC TTGCTCGCAA GCCCCGTGTA CTATCTGCAT CAGAACGACG TCATCTATGT GGAGCCGAAC GACAAGAAAA CACAGATGAG CGAGATCAAC CAGAATAATA ACGTAAACGT ATGGCTGAGT GTTACCTCCA CTTTGGTATC CATTCCACG CTGACGATTA CGATAATAGA TAAGACCAAA | 660 720 780 840 900 960 |
|----|--|--|
| 10 | (2) INFORMATION FOR SEQ ID NO:34 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1746 base pairs (B) TYPE: nucleic acid | |
| 15 | (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) NOLECULE TYPE: DNA (genomic) | |
| 20 | (iii) HYPOTHETICAL: NO (iv) AHTI-SENSE: NO | |
| 25 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc feature (B) LOCATION 11746 | |
| 30 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34 | |
| 35 | TCAAAAATAG TACTACGAAA GTTTTGTACC TTAGCGCGCA TGAAAAAAGAC CAATCTGTTT TTATCTCTGC TGGTGATCTT TATCACCGGT AGTTTTATGA CTGCCTGTGC ACAGAGTCC AAGACGAACA AACTCACCGA AGAAGATCGG AGCCGCAATG AGTATGTACA GTCGATGGAT GTGCTTAGCA ATATTATCGG TAACGTCAGG CTGTATTTCG TCGATACCAT AAGTATCAAA CATATGACTC GGCGTGGTAT AGATCGAATG ATTCCTTACG AGGAAATGGA TGAACTGAAA TTGATGACTA CGGGAGAGTA TGCCGGAGTC GGAGCTATCA TATCGCAGCG CCCGGATAGT GCTGTGATTA TCCAGAGACC TATGGAAGGT | 60 120 180 240 300 360 420 |
| 40 | ATGCCCGCAG ACGAAGCAGG ATTGATAGCA GGCGACCGCA TCCTGACTAT CGATGGGAAA GACTTCCGCA AATCCACCAC ACCGAAAGTA AGCCAAGCAC TGAAAGGGAT AGCCGGTACT GTTGCAAAGG TGACAGTAAT GCGCTATGGC GAAACCAAAC CTCGTACTT TTCCGTGAAA CGTCAAAAAAG TGATTATGAA TTCCGTCACT TACAGCGGAA TGCTCGATGG CTCGATAGGA TATATCCGCT TGAACAACTT TACGGACAAA AGTGCAGAAG AGGTGCGCAC GGCCTTGTTG GATCTTCGTG ACAAACAAGG AGCGAAAGGT CTCATTTTGG ATTTAAGAGG CAATGGTGGC | 480 540 600 660 720 780 |
| 45 | GGACTGATGC AGGCTGCTAT CGAGATAGTC AATCTGTTCG TCCCTAAGGG CAAAGAGGTG GTAACGACCA AAGGTCGCAT TGCAGAGTCG GCGTCCGTAT TTCGCACATT GACTGAACCG ATCGACACGA AACTCCCGAT AGTAGTCCTG ATCGATGGAC AATCGGCATC TTCCTCGGAG ATTGTAGCCG GAGCACTGCA GGATATGGAC AGGCCTGTAC TGATGGGACA AAAGAGCTAT GGCAAAGGGC TTGTACAAAC GACTCGTCAG CTACCATACA ACGGCGTGAT CAAATTGACT | 940 900 960 1020 1080 |
| 50 | ACGGCCAAGT ACTACATCCC AAGCGGACGT TGTATTCAGC GTTTGGACTA CAGCCGCACC AATCGGACAG GTATGGCAAC GGCCATTCCT GACAGTCTGC ACAAAATCTT TTACACTGCT GCCGGAAGAC GTGTAGAAGA TGCAGGAGGA ATCCTGCCTG ACATCGAGGT CAAACAAGAT ACGCTGCGA CATTACTTTA TTATATGGCC ATCAATAATG ACGTTTTCGA TTTCGTCACA GGTTATGTGC TCAAGCATAA AACGAATGGCC AAGCCGGAGG ATTTTTCCAT AACGAACGAG | 1140 1200 1260 1320 1380 |
| 55 | GACTATGCAG CTTTCTGCAA GATGATGGAA GAAAAGAAAT TTGACTATGA TCGCCAGAGT GGCAAGATGC TTGACAAACT GGAGGAACTG GCTAAGATAG AAGCCTACCT GCCGGAAGCC AACTCGGAC TTAAAAGCACT ACGCGAAAAG CTAAAACCCA ACCTGTCGCG TGATCTGCTA CACAATGAGA TTGTCACTCG CTATTATTAT GAGCGAGGCA GTATCCCCA GAGTTACCG GAAGATAAGC TAGTCAAAAA AGCTATTAAG | 1440 1500 1560 1620 1680 |
| 60 | CTGCTGAAGG ACCATCCGGA ACAAATTCGA CAGATCCTTG CAGCTCCGAA AGCAGAGAAT AAAGGG | 1740 1746 |
| 65 | (2) INFORMATION FOR SEQ ID NO:35 | 1746 |
| 70 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2955 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| - | (ii) MOLECULE TYPE: DNA (genomic) | |
| 75 | (iii) HYPOTHETICAL: NO | |

```
(iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 5
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...2955
10
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35
      GTAACAGACA AAATGCAAAA CAAAGGATTT GTGATTGTTA TCACATCGGC TCTGGCCATC
      ATCTGTGCGT TTTACCTGTC ATTCTCTTTC GTTACGAACC GTTACGAAAA GAAGGCTAAG
                                                                                  120
      GCGATGGGCG ATGTTGCCGG AATGGCCTAT CTTGATTCCA TGTCGAATGA GAAGGTCTGG
                                                                                  180
15
      TTCGGCTACA CGCTGAAAGA AGCTCAAGCC CAGCAAATTG GTCTTGGCCT TGACTTAAAG
                                                                                  240
      GGGGGTATGA ACGTTATCTT GAAACTTAAC GCAAGCGATC TGCTTCGTAA CCTCTCTAAC
                                                                                  300
      AAAAGTTTGG ATCCCAACTT CAACAAGCT CTGGAGAATG CTGCCAAGAG CACGGAGCAA
                                                                                  360
      TCCGACTTCA TCGATATTTT CGTGAAGGAA TATCGCAAGC TCGATCCCAA CGGTCGCTTG
                                                                                  420
      GCCGTTATCT TCGGTTCGGG TGACCTTCGC GACCAGATTA CCGCAAAGTC TACGGATGCA
                                                                                  480
20
      GACGTAGTGC GTCTGCTCAA AGAAAAATAT AATAGTGCTG TAGAAGCTTC GTTCAATGTG
                                                                                  540
      CTCCGTGCTC GTATCGATGC TTTCGGTGTG GTTGCACCTA ATTTGCAGCG ATTGGAAGGG
                                                                                  600
      CAAGGGCGTA TCCTTGTCGA ACTCCCCGGA GTGAAAGACC CTGAGCGTGT TCGTACCCTT
      TTGCAACGCA GTGCCAACCT ACAGTTCTGG CGTACATACA AATTCGAAGA GGTCAGCGGA
                                                                                  720
      GACTTGATCG CTGCCAATGA TCGTCTGAGC GAATTGGCTA TGAACAACAC GGATGCTACC
      CCGGAAACAG AGCCTGCAAC TACTGACTCT GTAGCTGCAA CAGCCGATTC TGCTGCTGTA CAAGCTGTAG CTGATTCTGC TACTGTAGCA CAAAAAGAGG CCAAGGATGC TACTCGTAAA
25
                                                                                  840
                                                                                  900
                                                                                  960
      GACGCACTCT TCTCTCTGCT TACTCCCGTG AATCGTGGCG GTGCAGTAGT GGGTGTGGCT
      CGTCGTGCTA ATATGGCTCA GATATCTGAA ATGCTCCAGC AAGCTCACGA TCTGAAGGTT
                                                                                 1020
      ACACGTGAAG ATGTGCTTTT CCTCTGGGGT GCTAAAGCAA TCGAAGACCC CGAAACCAAA
                                                                                 1080
30
      AAGGAGACCG ACCTCTACGA ACTCTATGCT ATTCGTACCA ATCGTACGGG AGATCCTGAT
                                                                                 1140
      TTGGGAGGTG ATGTAGTGAC TTCCGCCAAG AGTGATATCC AAAATGACTT CGGTCGTTCC
                                                                                 1200
      GAACCGATCG TTTCGATGAC GATGAATGAA GAAGGTGCTC GTAAATGGGC GCGTATCACA
                                                                                 1260
      AAGGATAACG TGGGACGGC AATCGCTATC GTTTTGGATG GTGTGGTTTA TTCTGCTCCG
                                                                                 1320
      AACGTGAATG ATGAGATCAC GGGCGGTCGC TCTCAGATCT CCGGGCACTT CACCGTGGAG
                                                                                 1380
35
      GAGGCCGGTG ACCTTGCCAA CGTACTCAAC TCCGGTAAAA TGGATGCTAC GGTAAGCATC
                                                                                 1440
      GAACAGGAAA ACGTGATTGG TCCTACGCTG GGTGCCGAGT CCATTAAAGC CGGATTCTTG
                                                                                 1500
      TCGTTCCTGC TCGCTTTGGT TATCCTGATG TGTTACATGT GTCTGGCTTA CGGTTTCTTG
                                                                                 1560
      CCGGGTCTTA TCGCAAACGG CGCATTGATT GTAAACAGCT TCTTCACATT GGGCGTATTG
                                                                                 1620
      GCTTCTTTCC ATGCCGTGCT GACCCTCTCG GGTATCGCAG GTTTGGTGCT GACGCTGGGT
                                                                                 1680
40
      ATGGCTGTGG ATGCCAACGT ACTTATCTTC GAGCGTATCA AAGAAGAGCT TCGTGCCGGT
                                                                                 1740
      AAGACTCCGA TTCGTGCCGT TACGGATGGT TATGGCAACG CTTTCTCTGC CATCTTCGAC
      TCGAACGTTA CGACTATTAT TACCGGTATC ATCCTATTCC TCTACGGGAC GGGGCCGATT
                                                                                 1860
      CGCGGTTTTG CCACTACGTT GATTATCGGT CTTATCGCTT CTTTCATTAC GGCTGTCTTC
      TTGACTCGTA TCGTCTTCGA GAAACTGGCG AAAAAAGGTC GTTTGGATAA GATTACATTC
                                                                                 1.980
45
      ACTACGAGCA TTACTCGCAA TCTCCTTGTC AATCCCTCAT ACAACATCTT GGGTAAGCGC
                                                                                 2040
      AAGACCGGCT TTATCATTCC GGTGATTATC ATCGTTTTGG GACTTATAGC TTCATTTACA
                                                                                 2100
      ATCGGTCTCA ATAGGGGTAT TGAATTCTCC GGAGGACGTA ACTACGTAGT TAAATTCGAC
                                                                                 2160
      CAGCCTGTAT CTTCCGAAGC CGTTCGTTCG GCCTTGTCTT CTCCCCTGCA GGAAAAGGTA
                                                                                 2220
      TTGGTTACCT CCATCGGTAC TGAAGGGACA GAGGTGCGTA TATCTACGAA CTATAAGATC
                                                                                 2280
50
      CAGGAGGAAA GCGAAGAAAC TGAAGCAGAG ATTACTGACA AATTGTATCA GAGCCTGAAA GGTTTCTACA CCCAGCAGCC TACTGCTGAT CAGTTCTTGG ACAATATCAT TAGCTCTCAG
                                                                                 2340
                                                                                 2400
      AAAGTAAGTC CCAGTATGTC GAGTGACATC ACGAGAGGTG CTATTTGGGC TGTGCTGTTA
TCGATGATCT TCATGGCCAT TTATATTCTG ATTCGCTTCC GTGACATTTC TTTCTCTGCC
                                                                                 2460
                                                                                 2520
                                                                                 2580
      GGGGTATTC3 TATCTGTGGC CGCTACTACA TTCTGCATTA TTGCTCTGTA TGCGTTGCTG
55
      TGGAAGATTC TGCCCTTCAC CATGGAGATC GATCAGAACT TCATCGCTGC TATTCTGGCT
                                                                                 2640
      ATCATCGGTT ACTCGCTCAA TGACACCGTG GTTGTATTTG ACCGTATCCG AGAGACGATG
                                                                                 2700
      AAATTGTACC CCAACAGAGA TCGCTATCAG GTGATCAACG ATGCCCTTAA TTCAACATTG
                                                                                 2760
      GGTCGAACAT TAAATACGTC TTTGACTACG TTTATCGTTA TGTTGGTAAT CTTCATCTTT
                                                                                 2820
      GGAGGTGCTA CGATGCGTAG TTTCACGTTC TCGATCCTGC TCGGTATCGT TATCGGTACA
                                                                                 2880
60
      TACTCTACGC TCTTTGTTGC TACACCCCTT GCCTACGAGA TCCAAAAGCG CAAGCTCAAC
                                                                                 2940
                                                                                 2955
      AAAGCAGCTA AGAAA
      (2) INFORMATION FOR SEQ ID NO:36
65
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 3138 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
70
                 (D) TOPOLOGY: circular
           (ii) HOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
75
```

(iv) ANTI-SENSE: NO

```
(vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
  5
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...3138
 10
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36
        TTTAAGTTTA ACGCAAAAGA AAAGTCTATG AAAAGAATGA CGCTATTCTT CCTTTGCTTG
                                                                                   120
        CTGACGAGCA TTGGGTGGGC TATGGCCCAG AATAGAACCG TGAAGGGTAC AGTTATCTCC
                                                                                   180
 15
       TCCGAGGATA ATGAGCCCCT GATCGGCGCG AATGTCGTGG TTGTCGGAAA CACCACTATC
                                                                                   240
        GGTGCTGCAA CCGACTTGGA TGGCAACTTC ACGCTTAGCG TGCCTGCCAA TGCCAAAATG
                                                                                   300
        TTGAGAGTGT CCTATTCCGG TATGACTACC AAAGAGGTCG CCATCGCTAA TGTGATGAAG
                                                                                   360
       ATCGTACTGG ATCCGGACTC TAAGGTTCTG GAGCAGGTAG TTGTATTGGG TTACGGTACG
GGACAGAAAC TCAGCACTGT TTCCGGTTCT GTGGCCAAAG TGTCCAGCGA AAAGCTCGCG
GAAAAGCCCG TTGCCAATAT CATGGATGCC CTCCAAGGTC AGGTAGCCGG TATGCAGGTT
                                                                                    420
                                                                                   480
20
                                                                                   540
       ATGACTACAT CCGGTGACCC TACTGCCGTC GCTTCTGTGG AGATCCATGG TACAGGGTCG
       TTGGGGGCAA GCTCTGCACC ATTGTATATC GTGGATGGTA TGCAAACTTC TTTGGATGTT
                                                                                   660
       GTGGCTACGA TGAATCCGAA TGATTTTGAA TCTATGTCCG TTTTGAAAGA TGCTTCTGCA
                                                                                   720
       ACATCTATTT ATGGAGCTCG TGCTGCAAAC GGAGTCGTTT TCATTCAAAC GAAGAAAGGT
                                                                                   780
25
       AAAATGAGCG AGAGAGGTCG TATTACCTTT AATGCCAGTT ACGGGATTTC TCAAATCCTG
                                                                                   840
       AATACTAAGC CCCTTGATAA TATGATGACT GGAGATGAAT TGCTGGATTT TCAGGTGAAG
                                                                                   900
       GCAGGTTTTT GGGGGAACAA TCAAACCGTT CAGAAGGTTA AAGATATGAT CCTTGCCGGA
GCTGAAGATT TGTATGGCAA TTATGATTCT TTGAAAGATG AGTATGGTAA GACATTGTTC
                                                                                   960
                                                                                  1020
       CCAGTGGATT TTAATCATGA TGCAGACTGG CTCAAGGCTT TGTTTAAAAC AGCACCCACC
                                                                                  1080
30
       AGTCAAGGTG ATATTTCTTT CTCCGGAGGG TCTCAGGGAA CTTCATATTA TGCCTCTATA
                                                                                  1140
       GGCTACTTCG ATCAGGAAGG TATGGCTCGT GAACCGGCAA ATTTTAAGCG CTATAGTGGC
                                                                                  1200
       CGGCTCAACT TCGAAAGTCG TATCAATGAA TGGCTGAAAG TTGGTGCAAA TTTGTCTGGT
       GCGATAGCGA ATAGACGATC TGCCGACTAT TTTGGAAAGT ATTATATGGG GTCAGGTACT
       TTCGGTGTGT TAACGATGCC TCGTTATTAT AACCCTTTTG ATGTGAATGG GGATTTAGCA
35
       GATGTCTATT ACATGTATGG AGCTACCAGA CCTTCTATGA CAGAACCGTA CTTCGCAAAA
                                                                                  1440
       ATGAGACCGT TCAGTTCCGA ATCACATCAG GCCAATGTAA ATGGTTTCGC CCAGATTACT
                                                                                  1500
       CCGATCAAAG GCCTTACTTT AAAGGCACAG GCTGGTGTTG ATATTACTAA TACTCGCACT
                                                                                  1560
       TCTTCTAAGA GAATGCCCAA TAATCCGTAT GATTCTACTC CTCTTGGGGA AAGAAGAGAA
                                                                                  1620
       AGAGCTTATC GAGATGTTAG CAAGTCTTTT ACAAATACGG CTGAATATAA GTTTTCAATT
                                                                                  1680
40
       GATGAAAAAC ATGATCTTAC AGCATTGATG GGGCATGAAT ATATTGAATA TGAAGGGGAT
                                                                                  1740
       GTTATTGGGG CATCTTCTAA AGGATTTGAA AGTGATAAGT TGATGTTACT GAGCCAGGGA
                                                                                  1800
       AAAACCGGAA ATAGTTTGTC TTTGCCTGAA CACAGAGTCG CTGAATATGC CTATTTGTCT
                                                                                  1860
       TTCTTTAGTC GTTTTAATTA CGGTTTTGAC AAATGGATGT ATATAGATTT CTCTGTTCGT
                                                                                  1920
       AATGACCAAT CCTCTCGATT CGGATCCAAT AATAGAAGCG CGTGGTTCTA TTCTGTCGGT
                                                                                  1980
45
       GGAATGTTTG ACATATAAA TAAATTCATT CAAGAAAGTA ATTGGCTCAG TGATCTTCGA
                                                                                  2040
       CTGAAAATGA GTTATGGTAC AACGGGTAAC TCGGAGATTG GTAATTACAA CCACCAAGCA
                                                                                  2100
       CTCGTTACTG TGAACAATTA TACTGAAGAT GCTATGGGGC TTAGCATTTC TACAGCAGGC AATCCCGACC TCTCGTGGGA AAAGCAGTCT CAGTTCAACT TCGGTTTGGC TGCAGGGGCT
                                                                                  2220
       TTCAATAATC GCTTATCTGC AGAGGTAGAT TTCTATGTCC GCACTACGAA TGATATGTTG
ATTGATGTCC CGATGCCTTA TATCAGTGGT TTCTTCTCAC AGTATCAGAA TGTAGGCTCT
                                                                                  2280
50
                                                                                  2340
       ATGAAAAATA CGGGTGTAGA CCTTTCTCTT AAGGGGACGA TCTACCAAAA TAAGGACTGG
                                                                                  2400
       AATGTATATG CTTCTGCGAA TTTCAACTAC AATAGACAGG AAATAACAAA GCTTTTCTTC
                                                                                  2460
       GGTCTCAATA AGTACATGTT GCCTAATACC GGTACTATAT GGGAAATTGG GTACCCCAAT
                                                                                  2520
       TCGTTCTATA TGGCTGAATA TGCTGGAATC GACAAAAAA CCGGTAAGCA GTTGTGGTAT
                                                                                  2580
55
       GTTCCTGGTC AAGTCGATGC GGATGGTAAT AAAGTTACAA CAAGCCAGTA CTCAGCTGAC
                                                                                  2640
       TTGGAGACAC GAATTGATAA GTCTGTTACT CCTCCTATTA CAGGTGGTTT CTCCTTAGGT
                                                                                  2700
       GCTTCTTGGA AAGGACTTTC TTTAGATGCT GATTTTGCCT ACATCGTTGG TAAATGGATG
                                                                                  2760
       ATCAATAATG ACCGTTACTT TACAGAGAAT GCAGGTGGAT TGATGCAATT AAATAAAGAT
                                                                                  2820
       AAAATGCTAT TGAATGCCTG GACAGAGGAT AATAAAGAAA CAGATGTTCC AAAATTGGGA
                                                                                  2880
60
       CAGTCTCCTC AGTTTGATAC GCATTTGTTG GAGAATGCTT CTTTCCTGCG TTTGAAGAAT
                                                                                  2940
       CTCAAACTCA CCTATGTACT CCCCAATAGT CTTTTTGCTG GGCAGAATGT GATTGGTGGA
                                                                                  3000
       GCTCGTGTCT ATTTGATGGC GCGCAATCTG TTAACTGTTA CGAAGTATAA AGGCTTTGAC
                                                                                  3060
       CCTGAAGCAG GGGGGAATGT GGGAAAAAAT CAATATCCTA ATTCTAAGCA GTACGTTGCG
                                                                                  3120
      GGTATTCAGT TGTCTTTC
                                                                                  3138
65
       (2) INFORMATION FOR SEQ ID NO:37
            (i) SEQUENCE CHARACTERISTICS:
70
                  (A) LENGTH: 2607 base pairs
                  (B) TYPE: nucleic acid
                  (C)
                     STRANDEDNESS: double
                  (D) TOPOLOGY: circular
75
           (ii) MOLECULE TYPE: DNA (genomic)
```

```
(iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
  5
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORTPHYROMONAS GINGIVALIS
           (ix) FEATURE:
10
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2607
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37
15
       TGGCATAGGA ATATTTTAT CTTTGCGAGT ACATTTAGCC CGAAAAATAT GCTCCCACTG
                                                                                  60
       CCATACCGTT ATGCAAAAAC CGAGCACCTT TTTCTCGCAA AAGGATACTG CAAGAATCCA
                                                                                 120
       ATAACAAACA TAATTATCCT ATTTATGAAG AAAAAGAATT TTTTGCTTCT TGGCATTTTC
                                                                                 180
       GTTGCTTTGC TGACTTTCAT CGGCAGCATG CAGGCACAAC AGGCCAAAGA TTATTTCAAC
TTTGACGAAC GGGGCGAGGC CTACTTCTCA TTCAAAGTGC CTGATAGGGC CGTTCTACAA
                                                                                 240
                                                                                 300
20
       GAGCTGGCTC TGATCATGTC CATCGACGAG TTTGACCCCG TAACCAATGA AGCCATTGCC
                                                                                 360
       TATGCCAGCG AAGAGGAGTT CGAGGCATTC CTGCGCTATG GGCTCAAGCC TACATTCTTG
                                                                                 420
       ACTCCTCCAT CCATGCAGCG CGCTGTCGAG ATGTTCGACT ACCGCTCAGG AGAAAAATAC
                                                                                 480
       GAATGGAATG CTTACCCCAC CTATGAAGCC TATATCAGCA TGATGGAAGA GTTCCAAACA
                                                                                 540
       AAGTATCCAT CACTTTGTAC TACTTCCGTC ATTGGCAAGT CCGTAAAGGA TCGTAAACTG
                                                                                 600
25
       ATGATTTGCA AGCTGACGTC CTCTGCCAAT ACAGGGAAAA AGCCTCGCGT GCTCTATACT
                                                                                 660
       TCTACGATGC ACGGAGACGA AACGACCGGA TATGTGGTAC TGCTCCGACT CALAGACCAT
                                                                                 720
       CTGCTGTCGA ACTACGAATC CGATCCGAGG ATTAAGAACA TTCTGGATAA AACGGAAGTA
                                                                                 780
       TGGATCTGCC CTTTGACCAA TCCGGACGGA GCATACAGAG CCGGAAACCA CACCGTACAA
                                                                                 840
       GGAGCTACTC GCTACAATGC CAACAATGTC GATTTGAACC GTAACTTCAA GGATGATGTA
                                                                                 900
30
       GCCGGTGATC ACCCCGATGG AAAACCTTGG CAGCCGGAGG CAACTGCATT CATGGATTTG
                                                                                 960
       GNAGGAAACA CCTCTTTCGT GCTCGGTGCC AATATACATG GAGGAACAGA GGTGGTGAAC
                                                                                1020
       TATCCATGGG ATAATAAAAA AGAAAGACAT GCAGACGATG AGTGGTACAA ACTGATCAGT
                                                                                1080
       CGCAACTACG CAGCCGCTTG TCAGAGTATT TCCGCCAGCT ACATGACCTC CGAAACCAAT
                                                                                1140
       TCGGGAATCA TCAACGGTTC AGACTGGTAT GTAATTCGCG GAAGTCGTCA GGACAATGCA
                                                                                1200
35
      AATTATTTCC ATCGTCTGCG AGAAATTACC CTTGAAATCA GCAACACGAA GTTGGTGCCG
GCCTCTCAAC TTCCAAAGTA TTGGAATCTG AACAAAGAAT CTCTGCTTGC TCTGATCGAA
                                                                                1260
                                                                               1320
       GAATCCTTAT ACGGCATCCA TGGTACAGTG ACTTCCGCTG CGAACGGACA GCCTCTCAAA
                                                                               1380
       TGCCAGATCT TGATAGAAAA CCATGACAAG CGCAACTCCG ATGTTTACTC CGATGCTACC
                                                                               1440
       ACAGGCTACT ACGTACGTCC TATCAAAGCC GGCACTTATA CGGTGAAATA CAAAGCCGAG
                                                                               1500
40
       GGTTATCCTG AGGCAACTCG TACCATTACG ATCAAGGACA AAGAAACCGT CATCATGGAC
                                                                               1560
       ATTGCATTGG GCAACTCGGT TCCTCTGCCT GTACCCGATT TCACAGCTTC TCCTATGACC
                                                                               1620
       ATCTCAGTAG GCGAAAGCGT CCAATTCCAA GATCAAACGA CAAATAACCC CACGAATTGG
                                                                               1680
       GAGTGGACGT TCGAAGGCGG ACAGCCTGCC ATGAGTACAG AGCAGAATCC GCTCGTATCC
                                                                               1740
       TATAGTCATC CCGGTCAGTA CGACGTTACG CTCAAAGTGT GGAATGCAAG TGGTTCCAAC
                                                                               1800
45
      ACGATTACGA AAGAAAAATT CATCACTGTC AATGCCGTTA TGCCTGTAGC TGAATTCGTC
                                                                               1860
      GGTACCCCGA CGGAAATAGA AGAGGGCCAG ACGGTATCTT TCCAAAACCA ATCCACCAAT
                                                                               1920
      GCCACCAACT ACGTATGGAT ATTCGATGGC GGCACTCCCG CTACCAGTGA AGACGAAAAC
                                                                               1980
      CCGACTGTGC TTTACAGCAA AGCCGGCCAA TACGATGTCA CGCTCAAGGC GATCAGTGCT
                                                                               2040
      TCCGGTGAAA CGGTGAAGAC GAAAGAAAAA TACATCACTG TCAAGAAAGC TCCGGTCCCT
                                                                               2100
50
      GCTCCGGTAG CCGACTTCGA AGGAACACCT CGAAAAGTAA AGAAAGGCGA GACAGTTACT
                                                                               2160
      TTCAAAGACT TGTCTACGAA CAATCCGACT TCATGGCTTT GGGTGTTCGA AGGCGGCTCT
                                                                               2220
      CCTGCCACCA GCACGGAGCA AAACCCGGTG GTCACCTACA ATGAAACAGG CAAGTACGAT
                                                                               2280
      GTCCAGCTGA CTGCCACCAA CGAGGGCGGA AGCAATGTGA AGAAAGCAGA AGACTACATT
                                                                               2340
      GAGGTTATCC TCGATGACAG TGTCGAGGAC ATAGTGGCAC AGACGGGTAT CGTCATTCGT
                                                                               2400
55
      CCGCAAAACG GAACGAAGCA GATCCTCATA GAAGCCAACG CTGCTATCAA AGCGATCGTT
                                                                               2460
      CTCTATGACA TCAATGGACG GGTCGTACTC AAAACTACTC CGAATCAGCT CCGCTCGACC
                                                                               2520
      GTAGATCTTT CCATCCTGCC CGAAGGAATC TACACCATCA ATATCAAAAC GGAAAAATCC
      GCTCGCACGG AAAAGATCCA TATCGGG
                                                                               2607
60
      (2) INFORMATION FOR SEQ ID NO:38
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 318 base pairs
65
                 (B) TYPE: nucleic acid
                 (C) STRAHDEDHESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
70
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
75
          (vi) ORIGINAL SOURCE:
```

```
(A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
 5
                  (B) LOCATION 1...318
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38
       CTTTCGTACA GTGGAGAGAG CGATGCAAAA GAGTCTGATC AGAATTGCCG GAAATGTACC
                                                                                    60
10
       TTCATCGGCT TTGAAAAACG AGTAAATACG ATGCGACTGA TCAAGGCTTT TCTCGTGCAA
                                                                                   120
       CTCTTACTGC TCCCCATTTT CTTCTACAAG CGGTTTATAT CGCCGCTTAC ACCGCCTTCA
                                                                                   180
       TGCCGGTTTA CCCCCTCATG TTCGTCCTAT GCCATCGAAG CCTTACGTAA ATATGGCCG
                                                                                   240
       GGCAAAGGAC TATTGCTGAG CATCAAGCGT ATTCTCCGCT GTCACCCGTG GGGTGGAAGT
                                                                                   300
       GGCTATGACC CCGTTCCG
                                                                                   318
15
       (2) INFORMATION FOR SEQ ID NO:39
             (i) SEQUENCE CHARACTERISTICS:
20
                  (A) LENGTH: 2583 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
25
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
30
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
35
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2583
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39
40
       TTAGAGATGG CATACGACTT TACACAAACA TTCCGCAACA GCCTGGAGTA CAGCTATCAG
      GAAGCAACCC GTCTCGGCGT CGTAGCCGTG ACGCAAGATA TGCTCGTACT CGGTATCATT CGCGACGAG ACAATGGCGC GATCGACATC ATGCGGCACT ATGGGATCAA CTTGTACGAA
                                                                                   120
                                                                                   180
       CTCAAACGGT TGATCGAGTT GGAAGCCATC GCCGAGAGTT TGCCTGCTTC GCCTGAGGGA
       TCGCCCATCT TCACCCCTTC GGCTCGGGAG GCTATCGATG ATGCCACAGA CATCTGTGCC
45
       GACATGGAGG ACGAGGCCGT CAGCCCGGTC CATCTGTTGC TGAGTATCCT CAACTCGACA
                                                                                   360
       CAGGAGAGCT TAGTACAAAA GATATTTATG AAACAAGGTA TAAAATACGA CACCATCCTG
                                                                                   420
       TCGGATTACT TCGGACAGCG CAACCCCTCC GAAGGGAAGT CTCCCTCCGA AATGGAGATC
                                                                                   480
       CTCGACGGGT ACCAAGACAA CGACTTCGAC GACGAAGAGG ACGAATCCTC TCCGCCTTCC
                                                                                   540
       GGGAATAGCG GGACAGGCGG AGGCTCCGGC GACGCCCCCG AACAGAATAC CGGCGGAGGC
                                                                                   600
50
      GATACTACCA CCACGACACG GAGTGGAGGC GACACGCCTG CACTGGACAC CTTCGGCACC
                                                                                   660
      GACATCACTG CCATGGCGGC AGCAGGCAAG CTCGACCCGG TAGTGGGTCG GGAGCAGGAG
                                                                                  720
      ATCGAAAGGG TGATACAGAT ACTCAGCCGG CGCAAAAAGA ACAATCCGGT GCTCATCGGC
                                                                                  780
      GAACCCGGTG TAGGCAAGAG TGCCATCGTG GAAGGACTGG CCGAACGCAT CGTGAACAGG
                                                                                  840
      AAGGTGAGCC GTATTCTTTT CGACAAGCGG ATCATCAGCC TCGATTTGGC TCAGATGGTA
                                                                                  900
55
      GCCGGCACCA AATATCGCGG ACAGTTCGAA GAGCGGTTGA AAGCCGTGCT CGATGAGCTG
                                                                                  960
      AAGAAGAATA CGCAGATCAT CCTCTTCATC GACGAGATAC ATACCATCGT GGGAGCAGGC
TCTGCAGCCG GATCGATGGA TACGGCCAAT ATGCTCAAAC CCGCTCTTGC
GTACAGTGCA TCGGAGCCAC TACGCTGGAT AGAACATAGA AAGGACCAGA
TACGCTGGAT AGAACATAGA AAAGGACCAGA
                                                                                 1080
                                                                                 1140
      GCACTCGAAC GCCGCTTCCA GAAGGTGCCG ATAGCCCCCT CGACTGCAGA AGAAACGCTG
                                                                                 1200
60
      ACCATCCTGC AAAACATCAA AGAGAAATAC GAGGACTATC ACGGTGTACG CTATACGGAC
                                                                                 1260
      GAAGCGATCA AAGCGGCAGT GGAACTGACC GATCGCTATG TATCCGATCG TTTCTTCCCA
                                                                                 1320
      GATAAGGCGA TAGATGCCAT GGACGAGGCC GGCGCGAGCG TCCATATCAC CAATGTGGTG
                                                                                 1380
      GCTCCGAAAG AAATCGAGAT ACTGGAGGCC GAATTGGCAT CGGTGCGAGA GAACAAGCTC
TCGGCCGTAA AGGCTCAGAA CTACGAACTG GCTGCCTCCT TCCGCGATCA GGAGCGGCGC
                                                                                 1440
                                                                                 1500
65
      ACTCAGCAGC AGATAGCGGA AGAGAAGAAA AAATGGGAAG AGCAGATGTC CAAGCACCGC
                                                                                 1560
      GAGACGGTGG ACGAGAATGT AGTGGCGCAT GTAGTGGCGT TGATGACAGG CGTTCCGGCT
                                                                                 1620
      GAGCGGCTGA GCACGGGCGA AGGCGAACGT CTGCGCACGA TGGCAGATGA TCTCAAGACC
      1680
                                                                                 1740
                                                                                 1800
70
      ATGATCAGGG TGGATATGAG CGAGTATATG GAGAAGTTCT CCGTTTCGCG TCTCGTGGGT
      GCCCCTCCGG GATATGTGGG CTATGAAGAA GGCGGCCAAC TGACGGAGCG CGTAAGACGC
                                                                                 1980
      AAACCCTATT CCGTGGTTCT CTTGGATGAG ATCGAAAAGG CGCATGCCGA TGTCTTCAAT
                                                                                 2040
      CTGCTCTTAC AGGTGATGGA CGAAGGTCAG CTGACCGACA GTCTGGGACG GCGCGTGAAT
                                                                                 2100
75
      TTCAAGAACA CCGTGATCAT CATCACCTCC AACGTGGGTA CACGCCAGCT CAAAGACTTC
                                                                                 2160
```

```
GGGCAGGGTA TCGGGTTCCG TTCGGAAAAA GACGAGGAAG CGAACAAGGA GCATAGCCGT
                                                                                    2220
        TCCGTGATCC AAAAAGCTCT GAACAAGACG TTCAGCCCCG AATTTCTCAA CCGTTTGGAC
                                                                                    2280
        GATATCATCC TCTTCGACCA ACTGGGCAAG ACGGAGATTC GCCGGATGGT GGACATAGAG
                                                                                    2340
        CTTAAAGCCG TCTTGGCGCG CATCCATCGT GCCGGATACG ACCTCGTCCT TACCGATGAA
                                                                                    2400
  5
        GCCAAGGATG TGATAGCGAC GAAGGGATAC GACCTCC*AT ACGGAGCACG ACCGCTCAAG
CGCACACTCC AGAACGAAGT GGAGGATCGC CTCACGGATC TTATCCTCTC CGGACAGATC
                                                                                    2460
                                                                                    2520
        GAGAAAGGGC AGACGCTTAC GCTCTCTGCT CGCGATGGCG AGATCATCGT ACAAGAACAA
                                                                                    2580
                                                                                    2583
 10
        (2) INFORMATION FOR SEQ ID NO:40
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1368 base pairs
 15
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
 20
           (iii) HYPOTHETICAL: NO
            (iv) AHTI-SENSE: NO
25
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROHOMAS GINGIVALIS
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
30
                   (B) LOCATION 1...1368
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40
       AGAGGAGGAC AAATCCGCCG ACACCATACC GATTCGTCAA GGGGATCGGA CAGCAAAGCC
                                                                                     60
35
       TGCCGCTTCT CCTTTCATGC CGAGACCATC GGTTTCTCCA ACCACCAAAG AGCAAAAATG
                                                                                    120
       AATTACCTGT ACATACTGAT TACACTTTTA CTCTCCGGCT TTTTTTCCGG TGCTGAGATT
                                                                                    180
       GCTTTCCTTT CTTCGGACAA ACTGCGTCTT GAGTTGGACA GGAATAGAGG CGATCTCACA
                                                                                    240
       GGCAGAGCGT TAAACTTGCT GTATCGACAT CCGGACCAAC TGGTGACTAC TCTCCTTGTG
                                                                                    300
       GGTAATAATA TCGTTTTGGT AGTCTATGGT CTGCTGATGG CGGGATTGCT GGCCGCACCT
                                                                                    360
40
       TTGGCGCAAT GGATTGATAA CGATGCTATG ATCGTCGTTC TCCAATCTGT CTTATCCACT
                                                                                    420
       ATCATCATAC TGTTTACCGG GGAATTTCTA CCCAAAGCCA TTTTCAAGAC CAATGCCAAT
                                                                                    480
       ATGATGATGA GGGTATTCGC CCTCCCTATC GTAGCGATCT ATTATCTGCT TTATCCTCTG
TCTAAACTCT TCACCGGTTT ATCTCGCTCT TTTATTCGTC TGGTGGACAA GAATTATGTG
                                                                                    540
                                                                                    600
       CCTACAACAG TAGGGTTGGG GCGCGTAGAT CTCGATCATT ATTTGGCAGA AAATATGTCC
                                                                                    660
45
       GGAGAAAACG AACAGAACGA CTTGACTACC GAAGTGAAAA TCATCCAGAA TGCGCTGGAT
                                                                                    720
       TTTTCCGGTA TTCAGGTGCG AGACTGCATG ATCCCACGCA ATGAGATGAT AGCATGTGAG
TTGCAAACGG ATATTGAAGT ACTCAAAACG ACTTTTATCG ATACCGGTTT GTCCAAGATC
                                                                                    780
                                                                                    840
       ATTATCTACA GACAGAACAT AGATGACGTA GTAGGATATA TCCATTCGAG CGAAATGTTT
                                                                                    900
       CGTGGGCAAG ACTGGCAAAA ACGTATCAAT ACTACTGTAT TCGTACCCGA AAGCATGTAT
                                                                                    960
50
       GCCAATAAAC TGATGCGACT ACTCATGCAG CGCAAGAAAA GCATTGCGAT CGTCATCGAT
                                                                                   1020
       GAACTTGGAG GTACGGCCGG AATGGTCACA TTAGAGGATT TGGTAGAAGA GATTTTCGGT
                                                                                   1080
       GACATTGAGG ACGAACACGA CACTCGCAAG ATCATAGCCA AACAGCTCGG CCCTCATACC
                                                                                   1140
       TATCTGGTCA GTGGTCGTAT GGAAATAGAT GATGTGAACG AACGTTTTGG GTTGTCCTTG
                                                                                   1200
       CCTGAGTCTG ACGACTACCT TACCGTGGCC GGATTTATCC TGAATAGCCA TCAAAATATC
                                                                                   1260
55
       CCACAGGCCA ATGAGGTCGT GGAGATTGCT CCTTATACTT TTACCATTCT CAGATCTTCT
                                                                                   1320
       TCCACCAAGA TCGAACTGGT GAAAATGTCC ATCGACGACC AATCGAAC
                                                                                   1368
       (2) INFORMATION FOR SEQ ID NO:41
60
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 897 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
65
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
70
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
75
```

```
(ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...897
  5
             (xi) SEQUENCE DESCRIPTION: SEO ID NO:41
        TTAAAAAAGG AGATAACTAT GAAACAGAAC TACTTCAAAA GAGTCTGCTC ACTGCTTTGG
                                                                                                60
        CTGGTTTTAC CCATGCTTAT TATGCCATTG GAAGTAGCAG CTCAAGAGAT TATTCCGAAC GAAGAGGTGT TGGAATCATT GACTTTCGTT GCACCGGTTG AGGAGACAGA CGCAATAGAG
                                                                                               180
 10
        GCAGAGGTAG AAGCTCTGCA GGAGATAGTC GCTACTGAGG AGATTGCGGA GCAGGCTGTT
                                                                                               240
        CGTTCTTATA CCTACACGGT CTATCGTGAT GGCGTGAAGA TTGCTTCAGG ATTGACTGAG
                                                                                               300
        CCCACTTTTC TCGATGAAGA TGTTCCTGCC GGCGAACATA CCTACTGCGT AGAAGTACAG
                                                                                               360
        TATCAGGGAG GCGTATCCGA CAAAGTATGC GTGGACGTAG AGGTGAAGGA CTTCAAACCG
                                                                                               420
        GTTACCAATC TCACCGGAAC TGCTTCCAAT GACGAAGTTT CTTTGGACTG GGACGGTGTG
                                                                                               480
15
        GAAGAGAAG CTGAAGAGCC GGCAAGTGAT AAAGCAGTCA GCTACAACGT CTACAAGAAT
                                                                                              540
        GGAACCTTGA TCGGTAATAC AGCTGAAACT CATTATGTGG AGACCGGTGT AGCCAATGGT
                                                                                               600
        ACATACATCT ACGAAGTGGA AGTAAAGTAT CCTGACGGTG TATCTCCGAA GGTGGCTGTA
                                                                                               660
        ACCGTGACCG TGACCAACAG CTCATTGAGC AATGTAGATG GACAGGCTCC TTACACATTG
                                                                                              720
        CGAGTAGAAG GCAAGAAGAT TATTGCGGAA GCCCATGGTA TGATCACGCT CTACGACATC AACGGACGTA CCGTGGCCGT AGCCCCGAAT CGATTGGAAT ACATGGCGCA AACCGGTTTC
                                                                                              780
20
                                                                                              840
        TATGCAGTGC GCTTCGATGT GGGGAATAAA CACCATGTAT CGAAAATACA AGTAAGA
        (2) INFORMATION FOR SEQ ID NO:42
25
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 1131 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
30
                    (D) TOPOLOGY: circular
             (ii) MOLECULE TYPE: DNA (genomic)
            (iii) HYPOTHETICAL: NO
35
             (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
40
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...1131
45
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42
        TTGTTTAATC ATAAAAAATC ATGGTATGAA ACATTTCAAT TTTATCTCGT TGTTTTCCGC
       TCTGGCTTTA TTCTTTTGTG TGAAAATACC CTTGCACAAC AAAAAACAGA GGAGTTTGCA
                                                                                              120
       CCTGTGTCGG ATTTACGTGC AGAAGCGTAC GGCTCTACCG TTTTCCTCCA CTGGACTCCG
                                                                                              180
50
       CCGTATGACA ATCCGATGAT TCCTCTAAGC GAGAGTTTTG AATCAGGTAT TCCAGCTATA
TGGAAGACCA TTGACGCAGA TGGCGATGGC TATAATTGGA TGCATTTGAC CAATTTCACG
                                                                                              240
                                                                                              300
       GGACAGAGTG GTCTCTGTGT CTCTTCGGCT TCATACATAG GCGGCGTCGG AGCTTTGACT
                                                                                              360
       CCGGACAATT ATCTGATAAC ACCCGAATTA AAACTACCCA CAGACGCGTT GGTGGAAATA
                                                                                              420
       ATCTATTGGG TATGTACTCA AGATCTCACT GCTCCATCGG AGCACTATGC CGTTTATTCC TCTTCTACAG GCAATAATGC TGCTGACTTT GTTAATCTCT TATATGAAGA GACTTTGACT
                                                                                              480
55
                                                                                              540
       GCCAAACGGA TACAATCCCC CGAGTTGATC CGCGGAAATC GGACACAAGG TGTTTGGTAT CAAAGAAAGG TGGTACTCCC TAACGATACT AAATATGTTG CTTTCCGCCA TTTTAATTCC
                                                                                              600
                                                                                              660
       ACGGATAATT TCTGGCTCAA TTTGGATGAA GTATCTATCC TGTATACCCC TCTTCCCCGA
AGAGCTCCGT GTCCGCATCC GGGTGGTTAC ACTTATTCTG TATTCCGTGA TGGACAAAAG
                                                                                              720
                                                                                              780
60
       ATAGCGAGTG GATTGTCGGC ATTGGCATAT ATCGATACGG ATGTACCGTA TGGGACTCAA
GACTATTGTG TCCAAGTCAA TTATCTGCAA GGAGACTCGT ATAAAGTCTG CAAAAATATA
                                                                                              840
                                                                                              900
       GTGGTGGCAA ATTCTGCAAA CATCTATGGG GCGGATAAGC CTTTTGCGTT GACCGTGGTT
GGCAAGACCA TTGTAGCGAG TGCTTTCAAA GGAGAGATCA CTCTTTATGA CATTCGTGGC
                                                                                              960
                                                                                            1020
       CGGCTGATAG CTTCCGGCTG CGATACGCTT AGGTACAAAG CGGAAAATGG TTTTTACCTC
                                                                                            1080
65
       ATTAAAATAC AGGTAAACGG AACTGTCTAT ACTGAGAAAA TCCAAATCCA A
                                                                                            1131
       (2) INFORMATION FOR SEQ ID NO:43
70
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 2547 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDHESS: double
                   (D) TOPOLOGY: circular
75
```

```
(ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
 5
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
10
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...2547
          (xi) SEQUENCE DESCRIPTION: SEQ TO NO:43
15
      TTCGGAATAT CACCTTCAAT GAAAAAAAGT TTTCTTTTAG CCATAGTAAT GCTCTTTGGC
                                                                             60
      ATTGCCATGC AGGGACATTC TGCTCCGGTT ACGAAAGAGC GAGCTTTGAG TCTGGCTCGG
                                                                            120
      CTGGCTTTGC GACAGGTATC CTTGCGAATG GGACAAACAG CAGTATCTGA CAAGATTTCC
                                                                            180
      240
20
      TCTCCTGCAT ATTTTTATGT AGCTAATCGT GGAAATAATG AGGGCTATGC TCTTGTAGCA
                                                                            300
      GCAGATGACA GAATACCGAC AATTTTAGCC TATTCACCCA TTGGCCGTTT CGACATGGAC
                                                                            360
      AGTATGCCGG ACAATCTTCG CATGTGGCTA CAAATTTACG ATCAGGAAAT AGGCCTGATA
                                                                            420
      CTTTCCGGAA AAGCTCAGCT CAATGAAGAG ATATTACGTA CCGAGGGCGT ACCGCTGAA
                                                                            480
      GTACATGCTC TGATGGATAA CGGTCATTTT GCCAACGATC CCATGCGATG GAATCAAGGT
                                                                            540
25
      TACCCATGGA ACAATAAGGA ACCACTGCTT CCTAATGGCA ATCATGCCTA TACCGGCTGT
      GTTGCTACTG CTGCAGCACA AATCATGCGC TACCATAGCT GGCCGCTTCA AGGTGAAGGC
                                                                            660
      TCTTTCGATT ATCATGCAGG TTCATTAGTT GGCAACTGGT CCGGCACATT TGGTGAAATG
                                                                            720
      TACGACTGGA TCAATATGCC CGGAAATCCC GACCTTGATA ATCTGACTCA ATCTCAAGTG
      GATGCCTACG CCACACTGAT GCGTGATGTG AGTGCATCTG TTTCGATGAG TTTTTATGAA
                                                                            840
30
      AATGGAAGTG GTACGTACAG CGTTTATGTA GTAGGAGCCT TGCGAAACAA CTTTCGCTAC
                                                                            900
      AAGCGTTCAC TGCAGCTACA TGTACGCGCC TTATATACCT CACAGGAGTG GCACGATATG
                                                                            960
      ATCCGCGGG AACTTGCCTC CGGAAGGCCG GTCTATTATG CAGGGAATAA CCAGAGCATA
                                                                           1020
      GGACATGCTT TCGTTTGCGA TGGTTATGCT TCGGATGGTA CTTTCCATTT CAACTGGGGT
                                                                           1080
      TGGGGAGGTG TTTCCAACGG CTTCTACAAA CTAACACTCC TCTCGCCGAC TTCGTTGGGT
                                                                           1140
35
      ATCGGAGGTG AGGGAATAGG TTTTACCATT TATCAAGAGA TCATCACCGG TATCGAACCG
                                                                           1200
      GCTAAGACTC CCGCTGAAGC CGGTACAGAT GCCTTGCCGA TCTTGGCACT GAAAGACATA
                                                                           1260
      GAAGCCGAGT ATAAAAGTGA ATCCGGATTG AACGTAGGGT ATTCGATATA TAATACAGGT
                                                                           1320
      GAAGAGCAAT CAAATCTTGA CCTCGGATAC AGATTGAACA AGGCTGACGG AGAAGTCATA
                                                                           1380
      GAGGTGAAAA CTTCATCTAT CAATATCTCT TGGTACGGAT ACGGAGAGCA TCCCGAGAGT
                                                                           1440
40
      TTCTCATTGG CACCTAATCA GTTGTCACAA GGAATCAACA CCATCACCCT ACTTTATCGT
                                                                           1500
      CGCACAGGCA CCGAACAGTG GGAGCCGGTA CGGCATGCAC AGGGAGGATA TGTCAATAGC
                                                                           1560
      ATTAAAGTAA ATACGACAGA CCCGAACAAT GTCGTAGTCA CGGTAGATAA TAACGAAGGC
                                                                           1620
      AAGCTCAGTA TCGTCCCCAA CAGCTTTGTC GCAGATCTGA ATTCTTATGA ACATAGTACG
                                                                           1680
      ATTACAGTAC AGTTCAATAG CGACAGCCCT GATGAGATCC GTACACCCGT AGCCTTTGCT
                                                                           1740
45
      CTATCTACAG GAGCTACTGC GGACGATGTA ATATCTTTGG GCTGGGTAAT GGCTGAAGTT
                                                                           1800
      CCGGGCGGTA GCAGCAACTA TCCGGTGGTT TGGTCTAAAG ACGTTCTCAC TCTCTCGGAA
                                                                           1860
      GGCGACTATA CATTGTGGTA TAGATTTTCC ATCAACAACC AAAAGGATGA ATGGAAAAAG
      ATCGGAAGCG TGTCAGTAAA AACACCGACA GAGTATACGC ACCCCTTATT CGAAGTGGGC
                                                                           1980
      CATANTCANA CTTCTACCTA TACGCTGGAT ATGGCACACA ACAGAGTATT GCCCGACTTT
                                                                           2040
50
      ACACTCAAAA ATCTCGGATT GCCTTTCAAT GGTGAGTTGG TTGTTGTTTT CCGCCAAACA
                                                                           2100
      CAATCCTCAT CGGGGTCTTT ATGGGCAGCT CAAGAAACAG TACATATCAA GCAAGGAGAA
                                                                           2160
      ACTITICGTAT ATAAACCTGT TGTCGAAGGC CCTATACCTG ATGGATCCTA TCGTGCGACC
                                                                           2220
      CTCCATGCAT TCGTAAACGG ACAACAACAG TTGTACCTCA AGGGGAAAAG GAACTACACG
                                                                           2280
      GTGAAGATCG TCAATGGTAC AGCGGTAGAA GCAATAGAAT CGTCAGAAGA GATCAGAGTA
                                                                           2340
55
                                                                           2400
      TTCCCTAATC CGGCACGCGA TTATGTGGAA ATATCGGCAC CTTGCATTCC CCAAGAAACA
      TCTATCATTC TTTTCGATCT GTCAGGCAAG ATTGTCATGA AGAATAGTTT ATCAGCGGGG
                                                                           2460
      CATGGCAGAA TGGATGTCAG CCGACTTCCT AATGGGGCCT ACATCCTTAA GGTGGATGGA
                                                                           2520
      TATACGACGA AAATAAATAT AGTGCAC
                                                                           2547
60
      (2) INFORMATION FOR SEQ ID NO: 44
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 885 base pairs
65
                   TYPE: nucleic acid
                (C) STRAHDEDHESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
70
         (iii) HYPOTHETICAL: NO
          (iv) AUTT-SENSE: NO
75
          (vi) ORIGINAL SOURCE:
```

```
(A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
  5
                  (B) LOCATION 1...885
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44
       ACGAAGAAAA CACTGATGAA AAAGCTATTT CTCTCGCTCA CGAGTCTTGT AATGGTCTTC
                                                                                   60
10
       GCTGTTGCAA GTTGCGATAT AATCGACAAG GATCAAACCC TCTTGCCGGC TCCGACCAAT
                                                                                  120
       GTGACACCCG ATAATCCGGA TGACAATCCT TCGGAGATCG ACATTACGCA GACGCACACA
                                                                                  180
       GAAAAATATG TTTTGGCTGA AGAATTTACC GGCCAAAAAT GTCTCAACTG TCCGAAAGGT
                                                                                  240
       CATCGCAAAC TGGCGGCTCT CAAGGAGCAA TACGGTAAGA GATTGACTGT TGTCGGTATA
       CATGCCGGCC CTGGATCTCT CGTGCCACCT CTTTTCCGTA CAGAAGCCGG AGACGCATAT
                                                                                  360
15
       TATAGCAAGT TCGCCAATAA TACCCCTCTC CCTGCGCTGA TGGTTTCGCG CAAAAAGTTC
                                                                                  420
       GGCTCTTCCT ACGTTTATGA TAAGAGCTAC AAAACGTGGG ACGTGCCTAT TGCCGAGCAG
                                                                                  480
       ATGGAGCAAA AGGCGAAGAT CAATATCTTT GCCGTGGCCG AATACACCGA TACCCAAAAG
                                                                                  540
       ATCAAGGTGA CTGTAAAGGG TAAAATACTG GAGGGGAATA CACTCCCGAA GTCCATGGTT
                                                                                  600
       CAGGTGTATC TGTTGGAGGA TAAGCTGATC GCTCCGCAGG TGGATGGCAA TACGACAGTC
                                                                                  660
20
       GAGAATTACG AGCACAATCA CGTGTTGCGT GGAGCCGTTA ATGGTATTTG GGGCGAAGAA
TTTGTGAATC TCAAAGATTA TTTGTATACT TACGCCGTTG AACCGCTCTC GGGTATGTCC
                                                                                  720
                                                                                  780
       TTCGTAGCCG AGAATTATTC GATTGTGGCT TTTGTATACG ATGTGCAGAC GTTCGAAGTG
                                                                                  840
       TATGACGTTG TGCATGTAAA GATCAATCCG CAATCCGATG GCAAA
                                                                                  885
25
       (2) INFORMATION FOR SEQ ID NO:45
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 690 base pairs
30
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
35
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
40
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
45
                  (B) LOCATION 1...690
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45
       ACGAATAAAA AAGAAGAGAC AATGAAGAAA TCAAGTGTAG TAGCCTCAGT TTTGGCCGTG
                                                                                   60
50
       GCTCTCGTGT TCGCCGGTTG CGGACTGAAC AATATGGCAA AAGGCGGCCT TATCGGCGCC
                                                                                  120
       GGAGTAGGAG GTGCCATTGG TGCCGGAGTA GGTAACGTAG CCGGAAATAC GGCTGTCGGT
                                                                                  180
       GCCATCGTCG GTACTGCAGT CGGTGGAGCA GCCGGTGCTC TCATCGGAAA GAAGATGGAC
                                                                                  240
       AAGCAGAAAA AAGAACTGGA GGCCGCAGTA CCCGATGCTA CGATTCAGAC AGTAAATGAC
                                                                                  300
      GGAGAGGCTA TTCTGGTTAC TTTCGATAGC GGTATCCTCT TTGCGACGAA CTCCAGCACT
CTGAGTCCCA ACTCACGCAC TGCGCTGACG AAGTTTGCTG CAAACATGAA CAAAAACCCC
                                                                                  360
55
                                                                                  420
       GACACGGATA TTCGTATCGT AGGCCATACG GACAATACCG GCTCCGACAA GATCAACGAT
                                                                                  480
       CCTCTGTCTG AGAGACGTGC AGCCAGCGTA TATTCTTTCC TGAATTCTCA GGGTGTGAGT
                                                                                  540
       ATGTCGCGCA TGGCAGCCGA AGGGCGTGGG AGCCATGAAC CGGTTGCAGA CAATAGCACA
                                                                                  600
       GTTGCCGGAC GTTCGGCCAA CCGCCGTGTG GAGGTTTATA TCTTGCCGAA TGCCAAGATG
                                                                                  660
60
       ATCGAACAAG CACAGCAAGG TACGCTGAAG
                                                                                  690
       (2) INFORMATION FOR SEQ ID NO:46
65
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1026 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
70
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
75
           (iv) AUTI-SENSE: NO
```

```
(vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 5
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...1026
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46
10
       AACAGGAACA GAAATATGTC GAAAAAATCG ATCCTTCTGC TTTGCTGTTC GCTGTGCTTC
                                                                                        60
       ATTTCTGCTA CGAAGGCTGT GACCCCCGTC AGAAATGTGC GCAATAGCCA AGTGAACAGC
                                                                                       120
       AAAGCAAAGA CCGAACGTAC AAAGCCCTCG GACTCTGTAC GGTACATTAG CAACATGATT
                                                                                       180
       GCAGATCGGC TGGAGTTCCG CAACAAGATT TCTTCCGAAA AAGAGGTAAG AAAAGCCGAA
TATGAAAATC GGCTGGCGAT GGAAGCACTC AATTACCCTG CCATAGATTT ATATGGTGAA
                                                                                       240
15
                                                                                       300
       GATTCTTGGA GCGAGTATGT AAACCCTTTC GTGGGTGCAG GAACCGATGT CGAAATTCCG
                                                                                       360
       AACTCCTATG ACATTGATTG CTCTTCGTTC GTGATGCCCG TCGAAGATAA GCAGGTCACC
                                                                                       420
       TCTCAATTTG GCTACCGTCG GCGTTTCGGA CGGATGCACT ATGGTATTGA TCTTTCAGTG
                                                                                       480
       AATCGTGGCG ATACGATACG AGCAGCCTTT GACGGGAAAG TTCGTGTACG CAGCTATGAA
                                                                                       540
20
       GCGCGTGGCT ATGGCTACTA CATAGTCTTG CGCCATCCGA ACGGACTGGA GACTGTGTAC
                                                                                       600
       GGACACATGA GTCGCCAATT GGTAGACGAG AATCAGATCG TTCGAGCAGG ACAACCGATC
GGATTAGGAG GCAGCACGGG TCGAAGCACC GGTCCTCATC TTCACTTCGA GACCCGCTTC
                                                                                       660
                                                                                       720
       ATGGGTATTC CCATCAATCC GAGTACCATT ATAGACTTCG ATAACGCAGT GCCGCTCCGA GACATTACA CATTCAAACG AGGGAGCAAT TCTCGCTATG CAAAAGCCTC TAAGACTTCT
                                                                                       780
                                                                                       840
25
       TCTCGCTATG CAAAAAAAGG GAAGAAAGGC AGACAAGCTT CTTCTCCTAT GACCTATAGA
                                                                                       900
       ATCAAAAAG GCGATACTTT GGAAACAATA GCCAAAAGGC ACGGCACTTC TGTTCAGAAA
                                                                                       960
       CTCTGTGCTA CCAATGGCAT TGGCAAGAGT AAAATTTTGA CTCCGGGCAA AGCCTTGAGG
                                                                                     1020
       ATCAAA
                                                                                     1026
30
       (2) INFORMATION FOR SEC ID NO: 47
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 477 base pairs
35
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                   (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
40
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SEUSE: NO
45
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc
                                       feature
50
                  (B) LOCATION 1...477
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47
       CCGAGCAAAA CGATAATTAA GACAATGGCA AAAATCAATT TCTATGCTGA AGGCGTCAGC
                                                                                       60
55
       CTTCCTCGGA TCAGAAGACG GATCGTCGGT AAGTGGATAG CCGAAGTATG CAGCCGATAT GGGAAAGCCG TGGGAGAAAT CTCCTATCTT TTCTGTGATG ACGAATATAT CCTGAAAGCC
                                                                                      120
                                                                                      180
       AATCAGGAAT TTCTCGATCA TGACTACTAC ACCGACATCA TCACCTTCGA TTCCTGCGAA
       GCGGATACGG TGAATGGCGA CCTGCTTATC AGTCTCGATA CCGTACGCTC GAATGCCCGT
                                                                                      300
       GCTCTTGATC TTCGATACGA AGACGAACTG CATCGTGTCA TTATCCACGG CATACTGCAT
                                                                                      360
60
       CTTTGCGGAT TGAAAGACAA GAGCAAAAAG GATGAAGCCC AAATGCGTGC AGCCGAAGAG
                                                                                      420
       AAAGCCCTTG TCATGCTGCG AGAAACCATC GGATCGGAGC TTTCCCTATT GCATACA
                                                                                      477
       (2) INFORMATION FOR SEQ ID NO:48
65
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1185 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRAHDEDNESS: double
70
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
75
```

```
(iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 5
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...1185
10
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48
       AAAAGTAAAA CTATGAAGGT AAAGTACTTA ATGCTCACAT TGGTTGGAGC AATTGCACTG
                                                                                         60
       AACGCAAGTG CACAGGAGAA TACTGTACCG GCAACGGGTC AGTTACCCGC TAAGAATGTT
GCTTTTGCTC GCAATAAAGC AGGCAGCAAT TGGTTTGTAA CACTGCAAGG CGGTGTTGCA
                                                                                        120
                                                                                        180
       GCGCAGTTCC TCAATGACAA CAACAACAAA GACCTCATGG ACCGCTTAGG AGCCATAGGT
15
                                                                                        240
       TCTCTTTCTG TCGGAAAGTA TCACAGCCCT TTCTTTGCAA CTCGTTTGCA AATTAACGGA
                                                                                        300
       GGTCAAGCCC ACACTTTCCT CGGAAAAAAT GGCGAACAAG AAATCAACAC CAATTTTGGT
                                                                                        360
       GCAGCTCACT TCGACTTTAT GTTTGATGTG GTTAACTACT TTGCACCATA TCGCGAAAAT
                                                                                        420
       CGTTTCTTCC ATTTAATTCC ATGGGTAGGT GTTGGCTACC AACACAAATT CATCGGTAGC
                                                                                        480
20
       GAATGGAGCA AAGACAATGT GGAATCACTG ACGGCGAATG TAGGAGTTAT GATGGCTTTC
                                                                                        540
       AGATTAGGAA AGCGAGTAGA CTTTGTGATC GAAGCACAAG CAGCTCACTC CAATCTCAAT
                                                                                        600
       CTAAGTCGCG CATACAATGC CAAGAAAACT CCCGTATTCG AAGATCCCGC AGGACGTTAT TACAATGGAT TCCAGGGGAT GGCTACAGCA GGTCTTAATT TCCGCCTGGG AGCCGTAGGC
                                                                                        660
                                                                                        720
       TTCAATGCCA TTGAHCCAAT GGACTACGCA CTTATCAATG ATCTGAATGG TCAGATTAAC
                                                                                        780
       CGTTTGCGCA GCGAGGTCGA AGAACTCTCA AAACGTCCTG TATCATGCCC CGAATGTCCT
25
                                                                                        840
       GAAGTAACTC CTGTTACTAA GACAGAAAAT ATACTGACGG AAAAAGCTGT ACTGTTCCGT
                                                                                        900
       TTCGACAGCC ACGTTGTGGA CAAAGATCAA TTGATCAACC TGTATGACGT AGCTCAGTTT
                                                                                        960
       GTAAAAGAAA CTAACGAGCC GATTACCGTT GTTGGTTATG CTGATCCTAC GGGTAATACT
CAATACAACG AGAAATTGTC TGAGCGTCGG GCTAAAGCCG TTGTTGATGT TCTGACAGGT
                                                                                       1020
                                                                                       1080
30
       AAATATGGTG TGCCTTCCGA ATTAATCTCT GTAGAATGGA AGGGCGACTC TACGCAACCG
                                                                                       1140
       TTCAGCAAGA AAGCTTGGAA TCGTGTTGTA ATCGTTCGCT CCAAG
                                                                                       1185
       (2) INFORMATION FOR SEQ ID NO:49
35
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1161 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
40
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
45
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROHONAS GINGIVALIS
50
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...1161
55
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49
       TATAAAATGA CATACAGAAT TATGAAAGCT AAATCTTTAT TATTAGCACT TGCGGGTCTC
       GCATGCACAT TCAGTGCAAC AGCCCAAGAA GCTACTACAC AGAACAAAGC AGGGATGCAC ACCGCATTCC AACGTGATAA GGCCTCCGAT CATTGGTTCA TTGACATTGC AGGTGGAGCA
                                                                                        180
60
       GGTATGGCTC TCTCGGGATG GAATAATGAT GTAGACTTTG TAGATCGTCT AAGTATCGTT
                                                                                        240
       CCTACTTTCG GTATCGGTAA ATGGCATGAG CCTTATTTCG GTACTCGTCT CCAATTCACA
                                                                                        300
       GGATTCGACA TCTATGGATT CCCGCAAGGG AGCAAGGASC GTAACCACAA TTACTTTGGA
                                                                                        360
       AACGCCCACC TTGACTTCAT GTTCGATCTG ACGAACTATT TCGGTGTATA CCGTCCCAAT CGTGTCTTCC ATATCATCCC ATGGGCAGGT ATAGGATTTG GTTATAAATT CCATAGCGAA
                                                                                        420
                                                                                        480
65
       AACGCCAATG GTGAAAAAGT AGGAAGTAAA GATGATATGA CCGGAACAGT TAATGTCGGT
                                                                                        540
       TTGATGCTGA AATTCCGCCT ATCAAGAGTC GTAGACTTCA ATATTGAAGG ACAAGCTTTT
                                                                                        600
       GCCGGAAAGA TGAACTTTAT CGGGACAAAG AGAGGAAAAG CAGACTTCCC TGTAATGGCT
                                                                                        660
       ACAGCAGGTC TAACGTTCAA CCTTGGCAAG ACAGAGTGGA CAGAAATTGT TCCTATGGAC
                                                                                        720
      TATGCTTTGG TCAATGACCT GAACAACCAA ATCAACTCAC TTCGCGGTCA AGTGGAAGAG
TTGAGCCGTC GTCCTGTTTC ATGCCCTGAA TGCCCTGAGC CTACACAGCC TACAGTTACT
                                                                                        780
70
                                                                                        840
       CGTGTAGTCG TTGACAATGT GGTTTACTTC CGTATCAATA GTGCAAAGAT TGATCGTAAT
                                                                                        900
       CAAGAATCA ATGTTTACAA TACAGCTGAA TATGCGAAGA CCAACAACGC ACCGATCAAG
                                                                                        960
       GTAGTAGGTT ACGCTGACGA AAAAACCGGT ACTGCGGCCT ATAACATGAA GCTTTCAGAG
                                                                                      1020
       CGTCGTGCAA AAGCGGTAGC CAAGATGCTT GAAAAGTATG GTGTTTCTGC GGATCGCATT
                                                                                      1080
75
       ACAATTGAAT GGAAGGGCTC ATCAGAGCAA ATCTATGAAG AGAACGCTTG GAATCGTATT
                                                                                      1140
```

WO 99/29870 PCT/AU98/01023

| | GTAGTAATGA CTGCAGCGGA A | 1161 |
|------------|---|---------------------------------------|
| _ | (2) INFORMATION FOR SEQ ID NO:50 | |
| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 585 base pairs | |
| 10 | (B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 15 | (iii) HYPOTHETICAL: NO | |
| 10 | (iv) ANTI-SENSE: NO | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1585</pre> | |
| 25 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50 | |
| 30 | GTAATTGGCA TCATTATGGA ATTTTTCATG TTATTCATAG CGGCGGTTTT CGTTAATAAC GTCGTGCTGT CGCAGTTCCT CGGTATATGC CCATTCTTAG GCGTATCGAA GAAGGTAGAC ACCTCAATCG GTATGGGTGC AGCCGTGACA TTCGTATTGG CACTGGCTAC CTTGGTTACC TTCCTGATTC AGAAGTTCGT TTTGGATCGT TTCGGATTGG GCTTTATGCA GACCATTGCA TTATTTTTGG TCATTGCCGC CTTGGTGCAG ATGGTGGAGA TCATACTCAA GAAAGTATCT CCTCCCCTCT ATCAGGCACT GGGTGTATTC TTGCCCTTGA TTACGACGAA CTGCTGTGTG | 60 120 180 240 300 360 |
| 35 | CTCGGTGTGG CTATTTTGGT TATCCAGAAG GATTATACCC TGCTCCAGAG CTTCGTCTAT GCAATATCCA CGGCTATCGG TTTCACCTTG GCAATGGTTA CTTTCGCAGG TATTCGAGAG CAACTCGATA TGACCAATCT CCCCAAAGCT ATGAAGGGAA TACCTTCGGC ACTCTTGGCT GCCGGTATAT TGGCTATGGC TTTCATGGGC TTCAGCGGTA TCGCC | 420 480 540 585 |
| 40 | (2) INFORMATION FOR SEQ ID NO:51 | |
| | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2628 base pairs | |
| | (B) TYPE: nucleic acid (C) STRANDEDNESS: double | |
| 45 | (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DMA (genomic) | |
| 50 | (iii) HYPOTHETICAL: NO | |
| | (iv) APTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 55 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12528 | |
| 60 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51 | |
| | TACCGATCTT ATCGTGGGAT AGGGAGTGGG ACACACTCTC CTAACCTCAA AAACCGACTA AAAAGGATCG GAATAAGGAT ACCGAACAGA CACTATATCC ATATCAAGCC AATCAAACCA | 60 120 |
| 65 | AAAAATAAAA TGAAACAACT AAACATTATC AGCTTCATCA TTGCTTTCCT ATTCTTAGGA ACGAGCGCAT CGGCTCAGCA ATCGGGCGGA TCCGTTACAG GTACCGTAGT GGACAAAAGC TCAAAAGAAC CTATCGCATA CGTACAAGTA TTCGTCAAAG GAACCACTCT CGGAACTTCC ACGGATGCAA ACGGAAACTA CTCGATCAAG GGAATCCCTT CGGGTAATCA AACTATCGTA GCCCGACTCA TGGGTTACTC CACTTGCGAA GAAAAAGTAC ATATAGAAAA GGGTGGTTCC | 180 240 300 360 420 |
| 70 | CGCCACGTAG ACCTCTATCT GACCGAAGAG ATTCTCTCTC TCGATGGGGT AGTGGTATCT GCCAATAGAAA ACGAGACTTT CCGCCGTCAA GCACCCTCGT TGGTAACGGT ACTGTCGCCG GAACTTTTCC TCAAAACCAA CTCTACCAAC CTGAGTCAGG GACTTAAGTT CCAGCCCGGT CTGCCGCTGG AGGACAACTG TCAGAACTGC GGTTTCAACC AAGTTCGTAT CAATGGACTC GAAGGAGGCCT ATTCGCAAAT TCTTATCGAC AGCCATCCCA TCTTCAGTTC GCTTGCCGGT | 480 540 600 660 |
| <i>7</i> 5 | GTCTATGGCT TGGAGCAGAT GCCTGCCAAT ATGATCGAAC GTGTAGAAGT AATTCGCGGT GGAGGTTCGG CTCTGATGCT GTGGGAGGCG TTATCAACGT AATTACGAAA | 720 780 840 |

| 5 | GAACCGCTTC GCAATTCGGC CGAGATCAGC CATTCTACGA TGACCTTCGA CCACGCGAAA GGGTGGGGGA GCTTCCAAAA TACGACCCAG TTCAACGGTT CTATGCTGAC GGAAGACCGC AAAGCCGGTG TCATGGTATT CGGCCAACAC AACTACCGTC CCGGACAGGA TATAGACGGC GACAACTTTA CCGAACTACC CAATCTGCGC AACCGCTCGC TCGGTTTCCG CTCATACTAT | 900 960 1020 1080 |
|----|---|--|
| | AAGACCGGTC TCTACAGCAA AGCAACCCTC GAATATCACA GCATGCAGGA GTACCGTCGT GGTGGCGACA GACTGGACAA TCCTCCTTTC GAAGCCCAGA TAGCGGAATA TCTCCAGCAC TATATCAATG GCGGAAGTTT CAAATTCGAT CAGGGCTTCA GCGGTGGCAA GGATTTCTTC AGTCTGTATG CTTCAGCACA AGACGTTCAG CGTCGTAGCT ACTACGGGGG TGGCGACTAT ACCGAAAATC TGCTGAACGG AGCAGTTCAG AGTGGAAGCA CCGAATCGGA CGAATACAAC | 1140 1200 1260 1320 1380 |
| 10 | GATGCTTTCA CGGCTCTTAC TTCCTACGGG ACTACCAAGG GATTCGATTT GCAAGGAGGA GGTATGTACC GCCATACCTT CGGAGAAAAC TGGGACTTTA CCGGCGGACT CGAATATATC TACGGCCCAAC TCGATGACAG AAGCGGCTAC AGACCGAGCA AAATAGATCA GAATACCTCT ACTTTTAGTC AGTACGACCA GCTCGAATAT AAGACGGAGA AGTTAAGTGC CCTTATCGGA GCACGTATCG ACTATGTTCT CCTCAATCAG GATGGCAAAAC GCTATATCGA TCCGCTCTTC | 1440 1500 1560 1620 |
| 15 | ATTTTCAGTC CTAGAGCCAA CGTACGATAC AATCCCAATA AGAATCTCAG CTTCCGACTC TCATACAGCG AAGGATTCCG CGCTCCTCAG TATTTCGATC AAGAATCTGCA CGTAGAGTTG GCCGGTGGTA CTCATCAG CCGTGTCCTT TCCCCCAATC TGAAAGAAGA ACGTTCACGA AGCATCAGTG CTTCTTCGA TTATTACCAC AGAGCCGACG AATGGCAATT CAATATCATG | 1680 1740 1800 1860 1920 |
| 20 | GGAGAAGCCT TCTCCACCTT TATCAGCAAT CAGTTCAAAC CATCCGATAA GGTCGAAACC ACGAGGATG GCAAAGAATG GATCATTCGT ACCATCTACA ACGACAAGGA TGGAGTATCG AAGGCTATTAG ACGACAAGGA TGGAGTATCG CTCGGCGGTA CATGCCAGAAGCCAAAACCACG GACAAGCCGA AAAACCAACG GACAAGCCGA GATCTCTGTG AAAGACTATG TACGCACTC GAATCTGTAC GGCTATTTCG TTGCTACGGT ACGTCCTACC GAGCACTCA CCATCAATCT CTCCGGTACA | 1980 2040 2100 2160 2220 2280 |
| 25 | TTCACGGGCA AAATGGATGT AGTACACGAA GCCTATGAAG GCGATATTCC CGCAGAACAC ATAGCTCCGG ACGGATCGTT CGACTTTGAA ATGAATGGTC AGCAATTCAA AGGTTTGGCC GAAGGTCATG CCAAGCTCGT CAAGACTCCG GCCTTCGCCG ATATAGACCT CAAGCTGAGC CACGACTTCC ACCTTGCTTC CACTATGACC TTGGAATTGA ATGCCGGAAT ACAGAACATA | 2340 2400 2460 2520 |
| 30 | TTCAACAGCT ATCAGAAAGA CACGGACAAG GGACCGGGTA GAGCTTCTAC TTACGTATAC GGTCCTATGC AGCCCAGAAG GATTTTCGTC GGTACAAAGA TCAATTTC | 2580 2628 |
| | (2) INFORMATION FOR SEQ ID NO:52 | |
| 35 | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 2697 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double | |
| 40 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 45 | (iv) ANTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| 50 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12697</pre> | |
| 55 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52 | |
| | ATATTGAATC ATTTGAGAAA GACTATGTAC AAAAAGATTA TTGCCGTAGC AGCTCTCTTC TGCGCCAGCA TAGGGATCCT GAAAGGACAG TCCTCGGATC TGACCCCTCA GGATACTATA TATAGCCCTG AAATATCCTA TGCCAAGCCT ATTCATAAGA CCATAGCATC TATTGAGATC | 60 120 180 |
| 60 | GAGGGAATGA GGTCTTTCGA TGACTTTGTC TTGCGCAATC TTTCAGGCTT GGCTGTAGGT | 240 |
| 00 | GATGAAGTCC TGATTCCTGG AGATGCCATG TCTGCTGCCG TGAATAGAAT TATGCGTCAG GGCTACTTCT CAAATGTGCG AATCATCGCG GATAAATATG TCGGCAATAA AGTCTATCTG | 300 360 |
| | AAAATCATTG TCACTGAACG TCCTCGCATC AGTAAGGTTA CTTTTAGCGG GGTAAAGAAG TCTGAGAGAG AAGATCTTGA AATGAAAATC GGTCTTCGCG AGGGGATTCA GATGACCAGA | 420 480 |
| 65 | AATAATGAAG ACAAGGTCAG GCAAATCGTA CAGAAGTATT TTAGTGAGAA AGGTTATCGC GATGCCAGCA TACGGATAAC GCAGGAACCG GATCTTTCCA AAGATGGCTT TGTCAATGTG | 540 |
| | CTTATCTCGA TTGAGAAGAA AAGCAAAACC AAGGTGAATG AAATTTATTT TTCCGGCAAC AAGGCCCTTA GCAATCATAA GCTAAGAATG GCGATGAAGA ACACCAATGC CAAATTCAGT CTTAGAAAGC ATATTCGCTC ATCTTTCTTG AAACCTTTTTA GTACTCATAA GTTTGTGGAA | 600 660 720 |
| 70 | GAGAGCTACC GTGAAGATTT GGTCCGATTG ATAGAGAAGT ATCAGGAATA TGGATATCGT GATGCTGAAA TACTGACCGA CAGTGTCGTG AAGGCTCCTG ACGGCAAAAG AGTGGATATT TATCTCAACA TCGAAGAGGG GCAGAAGTAT TATATTAAGG ATGTCAACTT TGTGGGCAAT TCACAATATC CATCGGAGTA TTTGGAACGA GTGCTCCGGAA TAAAATCCGG AGATGTGTAC | 780 840 900 960 1020 |
| 75 | AATCAGAGAC GATTGGCTAA GCGTCTCAAT GAAGATGAAG ATGCTGTGGG GAACCTGTAC TATAACAATG GCTATATTTT TGCGTGGGTC GATCCCGTGG AAACAAATGT AGTGGGGGAT TCTGTTTCGC TTGATATTCG TATAGCGGAG GGGAAGCAGG CCAATATCAA TAAGGTGATC | 1080 1140 1200 |

WO 99/29870 PCT/AU98/01023

```
ATCAAAGGAA ATACTGTCGT GTACGAAGAC GTAGTACGCC GAGAGCTTTA CACAAAGCCC GGCCAGCTCT TTAGTCGCGA GGATATCATT AACTCTATTC GTCTCATCAA TCAGCTTGGG
                                                                                     1320
       CATTTCGATG CCGAAAAATC TATTCCCCGT CCGATTCCCA ATCCCGAAAC AGGAACAGTG
                                                                                     1380
       GATATAGAGT ATGATTTGGT GCCGCGTAGC AGTGACCAAT TGGAGCTTTC TGTCGGTTGG
                                                                                     1440
  5
       AGTCAGTCCG GACTTCTGTT CCGAGGAGCC ATTAAGTTCA CGAACTTCTC TGTCGGCAAC
                                                                                     1500
       TTGCTCCATC CCTCGATGTA TAAGAAAGGG ATCATTCCGC AAGGGGATGG GCAAACACTA
                                                                                     1560
       TCACTGAGTG CTCAGACCAA TGGAAAGTAC TATCAGCAGT ATAGTGTCAC ATTTATGGAT
                                                                                     1620
       CCATGGTTTG GGGGCAAGCG GCCGGATATG TTCAGCTTCA GTGCATTCTA TTCCAAGACT
                                                                                     1680
       ACGGCGATTG ACTCCAAGTT CTACAATAGC AATGCCGGCA ACTACTATAA TGCCTACTAT
                                                                                     1740
       AATAGCTACT ACAACAACTA TAATAGTTAT TACAACGGTA TGTCGAACTA TACCGGCGAC
CTCTATACTC AGGCCAGCGA TCCGGATCGT TCGCTTCAGA TGTTAGGTAC TTCGATCGGT
10
                                                                                     1800
                                                                                     1860
       TACGGTAAGC GTTTGACTTG GCCGGACAAT TGGTTCCAGA TTTATACTTC TCTGAACTAC
                                                                                     1920
       ACCTACTATA GACTGCGAAA TTGGAGCTAC AATACCTTCC AAAATTTCCA TCATGGCTCG
                                                                                     1980
       GCTAATGATC TCAACTTGGA GCTGCGTCTC TCTCGTACTT CCATCGATAA TCCTATTTAT
                                                                                     2040
15
       ACCAGAAGCG GATCGGATTT CATGGTTTCT GTTGCTGCTA CTCTTCCTTA TTCTTTGTGG
                                                                                     2100
       GACAATCATG ACTATGCCAG CCAGAACCTC AGCGTAAGCG ATCGTTACAG ATTTATCGAG
                                                                                     2160
       TATCACAAGT GGAAGTTTAG AGGACGAGTT TTTACTCCAT TGCTCAATCC TGCTACGCAT
                                                                                     2220
       AAATATACAC CGGTGCTCAT GAGTCGAGTG GAAGGAGCAG TTCTTGGTTC GTATAATTCC
                                                                                    2280
       AATAAGAAAT CTCCTTTCGG TACTTTCTAT ATGGGAGGTG ATGGTATGTC CAGCTATTAT GGTGGCTACA TGAATGAGAC TATAGGTTTG CGTGGTTATA AGAACGGATC TATTGCCGGT
                                                                                    2340
20
                                                                                    2400
       AATAACTACG ACTATGCATA TGCTTATATG CGGCTTACGA TGGAACTACG TTTCCCGATT
                                                                                    2460
       CTGTTTGAAA ACTCATTCAA TGCGTGGCTC TTAGCTTTTG CCGAAGCAGG CAATGCGTGG
                                                                                    2520
       CGCAGTATCG ACAATTATAA TCCCTTTAAC CTGAAGCGAT CGGCCGGTGT AGGATTGCGT
                                                                                    2580
       GTAACGTTAC CGATGGTCGG AATGCTCGGT ATCGATTGGG GATATGGCTT TGACCGTCCG
                                                                                    2640
25
       GACAATTCTC TACAGCGAGG AGGAAGCAAT GTCCACTTTG TGCTCGGACA GGAGTTC
                                                                                    2697
       (2) INFORMATION FOR SEQ ID NO:53
30
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 531 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRAHDEDNESS: double
                  (D) TOPOLOGY: circular
35
            (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
40
           (iv) AHTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...531
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53
50
       CAAATAAACA TGAATGGCGA TATGAAACGG TTTTTGATTT TGATCGGCTT TGCACTGGCG GTAGCTTTCT CCGGTTTTTC CCAAAAGTTC GCTTTGGTAG ATATGGAATA TATCCTCAGG
                                                                                     120
       AATATTCCTG ACTATGAGAT GATGAACGAA CAGCTGGAAC AGGTGTCCAA GAAATGGCAA
                                                                                     180
       AATGAAATCG AAGCTCTCGA AAATGAAGCC CAATCTATGT ATAAGAAGTA TCAGAGCGAT
                                                                                     240
55
       CTCGTATTCT TGTCTGCTGC ACAGAAGAAA ACCCAAGAAG AGGCTATCGT AAAGAAAGAG
                                                                                     300
       CAGCAAGCAT CCGAGCTCAA GCGGAAGTAT TTCGGCCCGG AGGGGGAGCT GTATAAGAAA
                                                                                     360
       CGCTCCGATC TGATGAAGCC TATTCAGGAT GAGATTTGGA ATGCTATCAA AGAGATTGCC
                                                                                     420
       AAGCGTAACA ACTATCAGAT GGTGCTTGAT AGAGGTACGT CCGGAATTAT CTTTGCCAGT
                                                                                     480
       CCGTCTATTG ACATTAGCGA CCTTGTACTG AGCAAGATGG GCTTTAGCAA G
                                                                                     531
60
       (2) INFORMATION FOR SEQ ID NO:54
            (i) SEQUENCE CHARACTERISTICS:
65
                  (A) LENGTH: 510 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
70
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SENSE: NO
75
```



```
(vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
 5
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...510
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54
10
       CGAATAAATA AACAACACGA AATGAAGAAA TTTTTTCTCA TGCTTCTGAT GGCTCTTCCT
                                                                                       60
       TTGAGCCTCT TGGCACAAAA GGTGGCAGTG GTAAACACTG AGGAGATCAT TTCCAAAATG
                                                                                      120
       CCGGAACAAG TAGCTGCTAC CAAACAGCTC AACGAATTGG CCGAAAAGTA TCGCCTTGAT
                                                                                      180
       CTCAAGAGTA TGGACGATGA GTTTGCCAAA AAGACAGAAG AATTTGTAAA GGAAAAAGAC
                                                                                      240
       TCTCTACTGG AGAACATCCG CAATCGTCGT CAGCAGGAAC TTCAGGATAT TCAAACTCGT
                                                                                      300
       TATCAGCAGT CATACCAAAC GATGCAGGAG GATTTGCAAA AGCGCCAACA ACAGCTTTTT
GCTCCTATCC AACAAAAGGT GGCTGATGCC ATCAAGAAAG TGGGTGACGA AGAAAACTGT
15
                                                                                      360
                                                                                      420
       GCCTACATCA TGGAGGCCGG TATGATGCTT TACACCGGAG CTACTGCTAT TGACTTGACC
                                                                                      480
       GCAAAGGTAA AAGCGAAACT CGGAATCAAG
                                                                                      510
20
       (2) INFORMATION FOR SEQ ID NO:55
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2484 base pairs
25
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
30
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
35
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
40
                  (B) LOCATION 1...2484
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55
       ATCATGAAGG AAGCTATTCC CCGAAAGAAC AAGTATATAA AGCTCAACGG TATATACAGA
                                                                                       60
45
       TTGTCATTCA TTCTGCTATG CTGCCTGCTA TGCTCTCAGG CAGCTATGGC ACAAGGCGTC
                                                                                      120
       AGGGTATCGG GCTATGTGCT CGACCGTGGG GAAAAGCCGA TCCCGTTCGC CGGAGTCAAA
                                                                                      180
       GTGCGTGGTA CGGGGACAGG CGCAACGACG AATCTGAAAG GATACTACGA GTTTCGGATG
                                                                                      240
       AAGGCCACGA CGGACAGCAT CACGATCGAG TTCAGCTCCA TGGGGTACCA AGGGGTAAGT
                                                                                      300
       CGCAGCTTTC CGTCTCTGAC CAAGGACACT CGGCTGAATG TTCGTTTGGC AGAGGCCGAG
                                                                                      360
50
       ATGGAGCTTT CGAGCGTGAC GGTACAGGCC ACAAAACGCA GACTCAACAC GATGGAGCGC
                                                                                      420
       GTCAATACCC GAGACCTTCG TGTCAATGCA GGGCCTACGG GAGGGGTGGA ATCGCTCATC
                                                                                      480
       AGTACCTACG CAGGAGTAAC GCAGAACAAT GAACTAAGCT CGCAATACTC GGTTCGCGGA
                                                                                      540
       GGAAGCTACG ATGAGAATAT GGTCTATGTA AACGGAGTGG AGGTTTATCG CCCGCTGCTG
                                                                                      600
      GTTCGCTCTG CACAGCAGGA AGGTCTGAGC TTCGTCAATC CGGATCTGAC ACAATCCGTA CAGTTCTCCG CCGGAGGGTT CACGGCCGAC TATGGCGACA AGATGTCCTC CGTACTGGAT
                                                                                      660
55
                                                                                      720
      ATTCGCTACA AGCAACCGCA GGAGAAGGAA GGAGCGGTAC TCCTCGGGAT GCTACAATCG
                                                                                      780
      AGTGCCTACT ATGGCAGCAG TGCCGGAGCC TTCAGCCAAA TCACGGGTGT ACGCTACAAG
                                                                                      840
      AGTGCCAAAT CGCTCTTGGG CACTACGGAC ACGAAAGCCG AATACGATCC GATCTATGCG GACGGACAGA CATTCATGAC GTACCGTTTC AGCCCCAAGC TGTCGGTTAG TTTCCTCGGC
                                                                                      900
                                                                                      960
60
      AATATTTCGC AAACTCGCTA CAAGTTTGTC CCTCAGACCC GTGAGACGAG CTTCGGTACA
                                                                                    1020
      CTGAGCGATG CCAAAAAGTT GAAGATCTTT TTCGACGGTC AGGAACAAGA TCGTTTCCTG
                                                                                    1080
      ACCTACTTCG GTGCCTTCAG CATGAACTTC GTGCCGGACG ACAAACAGCG GCATACGGTT
                                                                                    1140
      ACGCTTTCGG CCTTCAACAG TAACGAACGG GAGACCTACG ATATTCAGGG AGAATACTTT
                                                                                    1200
       CTGAACGATG TGCAGCTGGG GGCGGACGGA ACTGCTTCGA TGGCTTCGGG CTCAGAGAAC
                                                                                    1260
65
       TCCAACGCT TGGGCATCGG GCGCAATCAC GAGCATGCGC GCAACAGGCT GAGCTACCGC
                                                                                    1320
      GTGCTGAACA TGGGTTACAG AGGGGAGATG AAGCTGAACG AGAAGCATCG CCTGCAAGCC
                                                                                    1380
      GGCGTATCGG CACAGATGGA GAAAATAGCC GACCATATCA GCGAATGGGA ACGGAGGGAT
TCGGTAGGAT ACAACCTACC TCACTCGGAG ACCGTATTGC TGATGTACAA TAACCTATAT
                                                                                    1500
      GCCGATACGC AGATGAGGGG AACGCGCTTG TCGGCATTCG TACAGGATCG ATTCAACTTC AGCATGGGAG GAGGTACATT TTCTCTCATT CCGGGTATCA GAGCTTCGTG GTGGAGCTTC
                                                                                    1560
70
                                                                                    1620
      AACAAGGAGT TGCTCGTCAG CCCACGTATC AGCGTGGGTT ATTCTCCCGA AAGCAACCCG
                                                                                    1680
      GCTTTGGTAC TGCGTGCAGC CGCCGGACTT TATTATCAGG CACCGTTTTA CAAAGAGCTA
                                                                                    1740
      AGGCAGACGC ATAAGGATGC CGAAGGCAAT AACGTGGTTG TCCTCAACGA GAAGATCCGC
                                                                                    1800
      TCTCAGGGAG CTTTTCACAT TCTCGCAGGA GCAGACTATA CCTTCGAAAT GGGGGGGCGA
                                                                                    1860
75
      AAATACAAGT TTACGGCAGA GGCTTACTAC AAGAGCCTGT TCAACATCAA CCCGTATATA
                                                                                    1920
```

| 5 | ATAGAGAACG TGAAGATCCG CTATCTGGGC GAAAACATCG GTTCGGGTTA TGCTGCGGGT ATCGATCTCA AGCTCTTCGG CGAACTGGTA CCCGGAGTGG ATTCGTGGCT GAAGCACCC ACTTACAAAT TCTCCTTCTT CCTTCAGGAG ACCGCTGCGGC CTGCACTAGA CGGAGGATTG CCCCAGCTCA ATCCGAGCC CATCACAGCC TCGCCGGCCT TTACCGCACC GGCCTATAAG CGTGTCGATC TGGGGGTTAT TGCGGCGGC ATGACTCTT TGCCGCCGA AGCAAATGG TAATGGGAGT AAAAGGGCC TACATAGGGG CTGACCTCTT CAATCTGTTC GACATGACA ACGTCAATTC TTACTACTGG GTGTCGGATG CCTACCAACA GCAATACGCC GTACCGAACT ACCTGACACG CCGCCAATTC AACCTGCGTC TCCTCGTCGA ATTC | 1980 2040 2100 2160 2220 2280 2340 2400 2460 2484 |
|----|---|--|
| | (2) INFORMATION FOR SEQ ID NO:56 | |
| 15 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2037 base pairs (B) TYPE: nucleic acid (C) STRAHDEDNESS: double (D) TOPOLOGY: circular | |
| 20 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 25 | (iv) ANTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 30 | (ix) FEATURE: | |
| 00 | (A) NAME/KEY: misc_feature (B) LOCATION 12037 | |
| 35 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56 | |
| 33 | CCTATCCCTT TTCTTCTTAA TATGTACAGC GGACATCATA AAATCCATTA TCCTTTCTT ATCCTGTTGG TATGCCTTGC TTTTGCTGCC TGCAAGAGCG TGAAGTTGAA AGATGCGGAG AAGGCACATG ATCGCCAAGA GTATACCAAG GCTGCCGATA TGTACAATAC ATTATACAGG CGTACCCGAC GAAAGCAGGT GGAGATGAAA GCTTATACGG CTTTCCGATC CGGTGAAAAAC | 60 120 180 240 |
| 40 | TATCGTGCCG CCGGCAGACA AGCCAAAGCT TTGCGTGGCT ATCTGAATGC CAGACGCTAC GGGTATCCGG ATTCTGTGGT ACTGCTCCGT TTGGCACAGA CTTATCAGCA AGGAGGTAAC TATAAAGGAAG CCGAGGTACT CTTCCGTGGA TATCTGGAAG CTTATCCGAA AAGTTATTTT | 300 360 420 480 |
| | GCAGCTATCG GTTTGGAGGG GTGTCTCTTT GCCCGCCAGC AAAAGGAATA TCCTACACGT TACCGGATAC GGCGAGCTGC CGAGTGGAAT TCGGCACGGG GCGACTTCGG CCCGGCCTAT | 540 |
| 45 | GCACCCGATG CTTCGGCTCT CTATTTCACA TCGAGCAGAA GCAAAGACGA CGGTTTGGAT AATAGCAGCA TAACGGGACT GAAACCCAAC GACATTTATA TCATCAAACG AGATGCACAA GGACGATGGG GACGTCCCGA TAGCGTGTCC GGAGGAATCA ACATCCATG GGATGAAGGC GTGCCAACGA TCACGCCCGA TGGTAGTACC ATATTATA CGTTGGCGCA GCAAGGAGCC GATTACGACC GTACGGTACA GATCTATTCC GCCGCTCGGA GCGAGAAGG CGGTTGGAGC | 600 660 720 780 840 |
| 50 | AACGGTTCGC TCGTGGACAT TATGCGCGAT TCGCTCCGTA TGGCTGCTCATG TCGCATCCG GCGATTACCT GTATTTCGTC AGCAATATAG GCGGTAGCTA TGGCGCAAG GATATTTATC GTGTCAAGGT GTCGGATCGT TCTTATGGTT CACCGGAGAA TTTGGGGCCT GATATCAATA CGCCGGGGGA CGAAATGTTT CCCTTCATAG ATGGGGATAG TACCCTTTTC | 900 960 1020 1080 |
| 55 | TTCGCTTCGG ACGGACACGC CGGTCTGGGA GGACTGGATA TTTTCAAAGC CACGCTGGAC TCTACCGGCC AATGGCATGT AGTCAATATG GGACAACCGG TCAATTCCTC TGCCGATGAT TTCGGCTTGG CTGTGGAGCC TAAAGGCAAA AACAAAGAAG AAGCTTTGCC GGACAACGGA GTCAAAGGTG TATTTTGTTC CAACCGAGGC GATGCACGG GTTGCACAC CCTCTTCCAT | 1140 1200 1260 1320 1380 |
| 60 | TTCGAACTGC CGGCTATCTA CACCGAGATT CAAGGTTATG TGATGGACAG AGAAGAAAAT CCCATAGCCG GAGCCACTGT CAGGATCGTA GGCGAACGCG GCCCCGTAGG ACAGGGATTC GTGACTACTC GTGACGATGG CTCCTATAAG ATGAGCGTGC AGGGCGATAC TCGCTATGTA ATGCTTGCCG GAGCATCGGG TTATTTGAAT CAGTACGTAG AACTCAAGAC CGATACCGCC | 1440 1500 1560 |
| | AAGCAGAGTG AGACCTACTA TGTGGACTTT TTCCTTGCAT CGCGTGAGAA AGCCGAGGGC TTGCAAAATA TTTTCTATGA TTTCGATAAA GCTACTCTTC GCCCCGAAAG CATGAAGAGC | 1620 1680 |
| 65 | TTGGACGAAC TGATTCGTAT CCTCACGGAC AATCCGGATA TTCGGATCGA ATTGGGTTCG CATGCCGACA GGAAAGGCCC CGATGCTTAC AACCTCGGAC TATCTGACCG CAGAGCCAAA TCCGTGGTG ATTACCTCAC GAGTCGTGCA ATAGCGCCG AAGGCTTAC GTGGAAAGGC TACGGTAAGT CTGTCCCCAA GACGGTGACA GCCAAAATTG CCGAACGGCA CGATTTCCTG AAGGAAGGCG ATGTGCTCAC CGAGGAATTC GTAGCACCTT TGACCGAGGA GCAGCAGTCA | 1740 1800 1860 1920 1980 |
| 70 | GTCTGCGACC AACTGAACCG TCGTACCGAG TTCCGTGTGA TCGAAGAAGA GTTGCGT | 2037 |
| | (2) INFORMATION FOR SEQ ID NO:57 | |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2316 base pairs

75



```
(B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
 5
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
10
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
15
                (A) NAME/KEY: misc feature
                 (B) LOCATION 1...2316
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57
20
      CCTGCACAGC CGAAAACATA TTGCATACGA TACTTTCGGA GAGAGGTATC TCCGAGGAGG
      CAAACAGAAA GGACATTGTA TGCGGACAAG ATGCGCCGGC ATATCGTCAA GTGCTGTCCC
      TCTGCCTCCA TACTCTTGCA AGAATACTCC ACAACAGAAT TATTCATTCA TCACTTTAAC
      ATATCAATAA TTATGAAAAA GTTTTTCTTC GCGCTACTAT CGATTGGTAT TTCAGCGCAG
      GCTTTTGCCA AGACGGACAA CGTCCCGACA GATTCGCTAC GAGTACACAA TCTTCAGACC
                                                                              300
25
      GTCACGGTCT ATTCTACACG CACGGCCGTA CCTCTGAAAA AGATACCGGC CAAGATGGAA
                                                                              360
      CTCATCTCAT CGCGCAACAT CAAGCAGTCC GGCTTTAACA ACATGACCGA CATCCTCAAG
                                                                              420
      ACGCAAAGTT CGCTCGATGT CATACAATAC CCGGGCTTTA GTTCGAACAT CGGTATCCGC
                                                                              480
      GGTTTCAAGC CCTCCGGCAA GTATGTAACC GTATTGGTAA ACGGCATCCC TGCGGGAACG
                                                                              540
      GACAATATCT CTACGCTCAA CACGAGCAAC ATCGAACAAA TCGAGATCCT CAAAGGCCCG
                                                                              600
30
      TTCTCTTCCA TCTACGGCAC CAATGCCATG GGCGGTGTGG TGAACATCAT CACCCACAAA
                                                                              660
      TCCAAGGACA AGATCCATGG CAACGTTTCT CTCTTCGGCG GTAGCTACCA GACCATGGCC
                                                                              720
      GGATCATTCA ACTTGGGTGG CCGCTTCGAG GATATTTTCT CATTCGATCT TAGTCTGGGC
                                                                             780
      TTGGACAAGC AGAACAAGGA CTATAAGACC GGATCAAACA ATTTCCTATC CCTGAGCAAA
                                                                              840
      CTGGAAGAAG CTATAGTAGA TGTAAATGCT ACCAAAAACA AGAAAATGAA GGGGAGCGAC
                                                                              900
      TATACTGTAG CAACGGGACG TCTGCGTTTC GGTATCGACT TCACGCCCGA ATGGTCGCTG
35
                                                                             960
      AATCTGTATC AAAACGTATT CCTCGGAGAT GCGATCCCCG TAGGAGGATC TATATGGGGC
                                                                            1020
      GTTTACGGAG AATCCAAAAA AAATCTGAAT CGTTCTTCGA CCTCTTTCGA GCTGCTCGGC
                                                                            1080
      AAACATGGCT GCCACACGCT TCAATTCTCC CCCTACTTCA ACATAGAGAA ATCGGAGAAC
                                                                            1140
      TATAACAATG CCGATCCCAC CGGTTTCATC AACTACAAAA GCGACTACTA CACCTATGGT
                                                                             1200
40
      GCCCTACTCC AGGACAAGAT TTCCTTTGGA GGACAAAATA TCGTACTCGG TGTCGACAGC
                                                                             1260
      CGAAACATGA CGATGGAGTC AGAAAGATTC GAGCAGGCAG GAGTGAATAC AAAGCCATAC
                                                                             1320
      AACCCCGGAT ATGCCACGAA CAATATCGGT TTGTTCGGAC AGGCCAATTT CTACCTGCTG
      AACGATGCTC TATCGATATC TGCCGGTGCA CGTGCCGACT TCATGTTCTT TGACCTGAAA
                                                                             1440
      GCGAACGAGT ATCTCAACAA TGAAGCCAAA CAGGAAACTC ATAACGTAAT CAATCCGAAT
                                                                             1500
45
      GTCGGAATCA AATATGAGTT TGTGAAAGGC CTTACAGCTC ATGGTACATT CGGTAGTGCA
      TTCAGTGCTC CCGATGCTTT CCAAAAAGCA GGCCAATACG TAGGCCCGTT CGGCACGACC
      ATAGGCAATC CTGACCTGAA ACCCGAAAAG TCCATGACCT GGGACTTCGG TATCGGATAC
                                                                            1680
      AGCAATGCAC GCTGCGGGAT CCAAGCCGAC GTAACCTTAA CCTATTTCCA CACCGACCAC
                                                                            1740
      AAAGATCTGA TCTTGTCCAG CCCTGACTAT GCTAATAATA TCACCACATA CATCAATGCC
                                                                            1800
50
      GACAAGGCTC GTATGAGCGG TATCGAGGCC CTTTTGTCTT ATGACTTCGG CAGCCTCTTT
                                                                            1860
      GCCAACAAGT TCTCTCCG CGCATTTGCG AATGCCACGA TCATGCTCAA TTCCGAGATG
                                                                            1920
      AAGAAAAGCC AGACCGATGC CCCTTGGAGC GAAATGTACT ACGTTCGCAA GCAGAACATC
                                                                            1980
      ACCTTCGGTA TCGAATATCG TGGCAAAGAA GGACTTGAAG TGATGCTCAA CGGTCGCTTC
                                                                            2040
      ATGGGACGCA GGATCGAGCA AAACTGGTAT GCTTACTACC CCGAAGTTCG CCCCGAACTC
                                                                            2100
55
      CAGCAACTGC TTGCAGCAGA AGAGCCTGAA TTGGCTGCTC AGGGACTGCT CCGTCATCCG
                                                                            2160
      CAAGCAATGG TGTTCAATGC CTCTGCTTAC TACCACATGA ACAAGTATCT CACCTTCGGT
                                                                            2220
      GTGAACTTGA ACAACATCTT GGATGAGCTT TATACGGAGA AAGACGGCTA CCACATGCCC
                                                                             2280
      GGACGTAACA TCATGGGTAA GGTTATGGTC AACTTC
                                                                            2316
60
      (2) INFORMATION FOR SEQ ID NO:58
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1452 base pairs
65
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
70
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
75
          (vi) ORIGINAL SOURCE:
```

```
(A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
 5
                  (B) LOCATION 1...1452
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58
       GGCCGGGCCT CTTCCCCTTA CAGGCAAATG GATGGCATAC TAAACGATGA ATATAGACAA
10
       GCATCTATGA ACAGGTTTTC AAATCATTGG CCCTGCATCC TCGTGGGGTT TGTACTCTGG
       TTTGTATCGG CGAGTCGGAC TGTGGCACAA AACGCCTCCG AAACGACGGT ATCGTACGAT
       ACGGATACCG CCGTACTCTC CGAAGCCGAT GTGCTTCGGA TCGCTCTTAG TGAGAATGCC
                                                                                   240
       ACAGTGAAAG TGGCCGATAT GGATGTGCGC AAACAGGAAT ATGCACGTAG GGCAGCACGT
                                                                                   300
       GCCGATCTCT TCCCGAAAGT AGACCTCAAT GGCGTTTACA GCCATACGCT AAAGAAGCAG
                                                                                   360
15
       GTCTTATATA TAGATATGCC CGGTTTCAGC AGTAGCGAAG GTATCGAAAT GGGGCGTACA
                                                                                   420
       CACAATACGC AAGGAGGGT GAACGTCTCC ATGCCATTGG TGTCGGCACA GCTTTGGAAA AGCATTGCCA TGACCGGAGA ACAGCTCGAT CTGGCTCTGG AGAAAGCTCG CAGCTCCCGA
                                                                                   480
                                                                                   5.40
       AGGATTGCCA TGACCGGAGA ACAGCICGAI CIGGCICIG ACTIONICAL ACTORITY ACCORDANCE TATTGTTGGC CGAGGACTCT TATGGCGTAT TCAAGCGCAG CTATGACAAT GCTCTGGCCA ATTATAAGAA CATATCCGAC
                                                                                   600
                                                                                   660
20
      AAGTTCGATC GTGGACTTGT GGCCGAGTAT GATAAGATTC GAGCCAATGT ACAGGTACGC
AACATCGAGC CTAACCTCTT GCAAGCGCAG AACTCCGTAG CCCTTGCTCT CTGGCAGCTC
                                                                                   720
                                                                                   780
       AAGGTCCTGA TGAGCATGGA AGTGGAAACT CCGATCAGAC TCTCCGGTTC ATTGTCCGAC
       TATAAAGAAC AAGTCTATAC CGGCTATTTT GCCGCCGATA CGCTTATTTC CAACAACTCC
                                                                                   900
       TCCCTGCGTC AGCTCGATAT ACAGCGTCGT CTGGCTGTCA GTGCAGACAA GCTGAACAAG
                                                                                   960
25
       TACAGCTTCC TGCCTACACT CAATCTGGGA GGGCAGTACA CCTATTCGCT CAACAGCAAC
                                                                                  1020
       GACATCAAAT TCTGGGGCGA GGGACAACGC TGGACGCCTT TCTCCACCAT ATCGCTCAGC
                                                                                  1080
       CTGTACATTC CTATATTCAA TGGAGGCAAA CGTCTGTACA ACGTGAAGCA AAGTGCTTTA
                                                                                  1140
       TCGATCCGTC AGATCGATCT GCAACGACGC CACATAGAGC AATCCATCCG AATGGGAATC
                                                                                  1200
       AAGAACCAAA ATGACCGTCT GCGTACCTGT ATGCAGAGAT TTGTGGCCTC GGAAGAGGCT
                                                                                  1260
30
       GTCCGAAGTG CAGAAAAGGG CTATCAGATA GCAGAGAAAC GCTATCAGAC AGGCGAAGGC
                                                                                  1320
       ACTCTCGTCG AGCTCAACGA TGCCGATGTG GCTCTTTTGC AGGCTCGACT CAATTATAAT
                                                                                  1380
       CAGGCCATAT TCGACTTTAT GACCGCAAAG GCCGAATTGG ACAAGATGAA CGGCATGGGG
                                                                                  1440
       ATTCCCGAAC AA
35
       (2) INFORMATION FOR SEQ ID NO:59
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1620 base pairs
40
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
45
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SENSE: NO
50
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
55
                 (B) LOCATION 1...1620
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59
      TTTCATAACT TTGACTTCCT AAACGGTATA AAATTGTTTT CGATGGCAAA TAATACTCTT
60
      TTGGCGAAGA CTCGACGTTA TGTCTGCCTT GTCGGTTTCT GTTGGCTCAT GGCGATGATG
                                                                                   120
      CACCTCTCTG GGCAGGAAGT CACTATGTGG GGGGACAGCC ATGGAGTGGC GCCGAACCAA
      GTGCGCCGAA CGCTGGTGAA GGTAGCCTTA AGTGAATCCC TTCCTCCGGG TGCAAAACAG
                                                                                   240
      ATTCGTATCG GATTCTCTCT TCCGAAAGAA ACGGAGGAAA AAGTCACCGC CCTATATCTC
      CTTGTGAGTG ATTCTTTAGC GGTGCGCGAC TTGCCGGACT ACAAAGGGCG AGTCTCTTAC
                                                                                   360
65
      GATAGCTTCC CGATCTCAAA GGAAGATCGT ACCACAGCCC TTTCTGCGGA TTCGGTAGCC
                                                                                   420
      GGACGCCGCT TCTTTTATTT GGCTGCGGAT ATAGGGCCTG TTGCTTCTTT TTCCCGATCC
                                                                                   480
       GATACGCTGA CTGCCCGTGT GGAAGAGGTG GCTGTCGATG GCCGCCCTTT GCCGTTGAAA
                                                                                   540
      GAGCTGTCGC CTGCCTCCCG TCGTCTGTAT AGGGGGTATG AGGCCCTCTT TGTACCCGGT
                                                                                   600
      GATGGCGGAT CGCGGAACTA TCGTATCCCG GCCATTTTGA AAACGGCTAA TGGAACACTC
                                                                                   660
70
      ATAGCGATGG CCGACAGACG AAAATATAAT CAGACGGATC TGCCGGAGGA TATAGATATA
                                                                                   720
      GTCATGCGGC GCAGTACGGA CGGAGGGAAA TCGTGGAGCG ATCCCAGGAT TATCGTACAG
                                                                                   780
      GGAGAGGGGC GCAATCATGG CTTTGGCGAT GTAGCCCTGG TGCAAACCCA AGCAGGAAAG
                                                                                   840
      CTCCTGATGA TCTTTGTCGG TGGAGTAGGC CTGTGGCAGT CTACCCCCGA TCGTCCTCAG
                                                                                   900
      CGCACTTATA TATCGGAAAG TCGGGACGAA GGACTGACTT GGTCGCCTCC TCGGGATATA
                                                                                   960
75
      ACCCATTTCA TCTTCGGCAA GGATTGTGCC GATCCGGGAC GCAGTCGCTG GTTGGCCTCC
                                                                                 1020
```

```
TTTTGTGCTT CGGGACAAGG GCTTGTGCTG CCATCCGGTC GTATCACGTT TGTGGCTGCC
                                                                                   1080
       ATCCGCGAAT CAGGGCAGGA GTACGTCCTG AACAACTATG TCCTCTATAG CGACGATGAG
                                                                                   1140
       GGCGATACAT GGCAGCTTTC CGACTGTGCA TACCGCCGTG GCGATGAGGC AAAGCTTTCA
                                                                                   1200
       TTGATGCCCG ATGGCAGGT ACTGATGAGC ATACGCAATC AGGGACGGCA GGAGAGCCGA
                                                                                   1260
 5
       CAGCGTTTCT TCGCTCTCTC CTCCGACGAT GGCCTTACTT GGGAGAGAGC CAAGCAGTTC
                                                                                   1320
       GAGGGCATCC ATGACCCCGG CTGTAATGGA GCTATGCTTC AAGTGAAAAG GAACGGAAGG
                                                                                   1380
       GATCAAGTGC TGCACTCCCT GCCTCTCGGC CCGGATGGGC GTCGCGATGG AGCTGTCTAT
                                                                                   1440
       CTCTTCGATC ATGTCTCCGG CCGCTGGTCC GCTCCCGTTG TTGTCAATTC AGGATCGAGT
                                                                                   1500
       GCCTACTCGG ATATGACTCT GCTGGCGGAT GGAACGATCG GTTATTTCGT CGAAGAGGGC
                                                                                   1560
10
       GATGAGATCT CATTGGTTTT CATTCGGTTC GTCCTTGACG ATCTCTTCGA TGTCCGGCAA
                                                                                   1620
       (2) INFORMATION FOR SEQ ID NO:60
15
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 879 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
20
            (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
25
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
30
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...879
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60
35
       AAGTCTCCGA GCGATTCGGC TTTGCCGTCG GTATGGAGCG CGAATACAAT ATCTGGACTC GTCGGTGGGA AACGCATTAC TTTGCTTATC CTGTATTCTA TGGCGATAAG AAGTAGTAAT
                                                                                     60
                                                                                    120
       ATAGAATCAA TACAGTGCTT TGTGATGAAA AAAGAAAAAC TTTGGATTGC GATCGTCGCC GGTTTGGCTT TCGTATTGGG CCTTTATGCT CTTGGCCGCA GTGTCGCTCA GCTACGCCGC
                                                                                    180
                                                                                    240
40
       TCTCAGCCTT CGGTGACTGT GACCGGTATG GCCGAGCGTA ATTTCAAATC CGATCTGATC
                                                                                    300
       GTTTGGACTG CTTCGTACCA GCTCCAGATG ATGGATCTCG AATCGGCCTA CAAGGCTTTG
                                                                                    360
       AAGGAAAAAC AGATATTGGT AGCAGACTAT TTGAAAAAACA AGCAGCTGCC CGATTCGTCT
                                                                                    420
       TATATCTTCT CAAGCGTAGC CATCTCTAAA GAATACAACT ACTATTACGA TCCTCGGCAG
                                                                                    480
       GAACAAAACG TCAGGACCTT TGCCGGGTAT CTGCTCAGCC AGACAGTTAC GGTGACCTCA
                                                                                    540
45
       CAGGACATCG AACATGTGGA GAAAATATCT CGCGATATAA CGGAGCTGAT CAATCAGGGG
                                                                                    600
       GTAGAGATTA CCTCCGACCG TCCGGCCTAT TACTACACCA AGCTCAATGA TCTGAAGGTG
GAGATGCTGC GCAATGCCTC CGAAGACGCT TTCAATCGTG CTTCGGTCAT TGCGGAGGGG
                                                                                    660
                                                                                    720
       AGCGGTTCCT CCGTGGGTAA GATGCTATCT TCTTCGATGG GCGTGTTCCA GATAGTGGGG
                                                                                    780
       CTCAACTCGA ACGAAGATTA TAGCTGGGGA GGTTCGTTCA ATACGTCTTC CAAGATGAAG
                                                                                    840
50
       ACGGCAAGCA TAACGGTTAA GGCTTCTTTC GCTTTGAAG
                                                                                    879
       (2) INFORMATION FOR SEQ ID NO:61
55
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 840 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
60
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
65
           (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
70
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...840
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61
75
```

```
GGGAAACTCC AAATGAAAAA AACAATTGCA ATTATCGCCT CAGCCCTCTT GGCTTTAGGA
                                                                                      60
       GCCGTCGGCT GTAAGAAAAA TGCTGACACT ACCGCTGTCA GTGAAAAGGA TAGCATAGCC
                                                                                     120
       TTGTCCATGG GTATTTTGTA CGGACAGGAT TTTGCCAATC AGTTCGAAAT GTCCCGCTTG
                                                                                     180
       CAAGGCCAGC CGATTGATTC GGTAGCTTTC TTGGACGGTT TCAAATATGG TATCGATACG
                                                                                     240
  5
       ACGCGCTTCT CGTACAATCT GGGAGCCATC TATGCTTCCA ATATAGCTCG TCAGCTGGCT
                                                                                     300
       CATGATTCCA TCGATATCGA CAAGTTCTAT GCAGCCATGC GTGCGGCTCT TCTTAAAGAC
                                                                                     360
       ACCGTATCTA TCGCCATGAA GCCTGCAGAT GCACAGGCTT TCATGCAACG AATCCAAGCC
                                                                                     420
       AAAAAGCAGC GAGAAAACAA TATGAAGCAG TTTGGCCAGA ACATCGAAAA GGGTAATGAA
                                                                                     480
       TACATCGATA CCTTTAAAAA AGAAGATGGT GTAACTGTTA CGACAACTGG TCTGGCATAC
                                                                                     540
10
       AAGACTCTTC AGGAAGGTAC GGGAGCTACT CCCTCTTTGG CCGATACTGT ACGTGTCAAG
                                                                                     600
       TATGTGGGTA CTCTGGTCGA TGGTAAAGAG TTCGACAAAA ACGAAGAAGG AATCGAATTT
                                                                                     660
       GCCGTTACCG GTGTGATTAA AGGCTGGACG GAGATGCTCC AACTCATGAA GGTCGGTCAG
                                                                                     720
       AAAGTTCGCG TGGTAATCCC ACAGGAGCTG GCTTATGGGG AGACCGGCAA CTATACCATC
                                                                                     780
       GAACCGTTCT CTACCCTGAC GTTCGAGATG GAACTTATCG GGATCAAGCC CGGGAAAAAG
                                                                                     840
15
       (2) INFORMATION FOR SEQ ID NO:62
             (i) SEQUENCE CHARACTERISTICS:
20
                  (A) LENGTH: 2409 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
25
           (ii) HOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
30
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
35
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2409
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62 '
40
       TGTCGAAAGC AGAAGCTCTC AAACCCGAAG AGGAGCCGGT ATCCGTGCAG ACGGATATCA
       TTCCGACAAA GCGATAAGAA TCCGATGAAA GTATTACGGC AAGTATTCCT CCCCATCCTT
                                                                                    120
       TTTGTCCTAC TGACAGGTGC CTGCTCCACC ACAAAGAATC TGCCGGAAGG CGAACAGCTG
       TATATCGGAA TGGGCAAGAC ACAGATACTC CGGCAGGACA AGAGCCACGC CGGCCAACAG
                                                                                     240
       GCTCTGACCG AAGTGGAGAG TACACTGAAA GTTACACCCA ATGGAGCTAT TTTCGGCAGT
                                                                                     300
45
       GCAAGTGCCT CCTTACCCAA GATACCATTC GGGCTATGGC TATACAACAG CTTCGTGGGG
                                                                                     360
       GATTCCACTG TCATTTCGAA ATGGATATTC GACAAGTTTG CAGCCAAGCC GGTTTTCATC
                                                                                     420
       AGTCAGGTCA AATCCGATAG CCGGGCTAAG GTGGCGACGA ACATCCTCCG CGAACACGGG
                                                                                     480
       TACTTCGATG CTAAAGTAAA AAGCAGTGTG ACCACTCTGA AAAAGGACTC GCTCAAAGCC
                                                                                    540
       AAAATCTCCT ATACGGTGGA TATGGCCTCT CCTTATCATT ACGACAGCAT CATTCCCTTA
CCGATCAGCA CTTTCCCCGA CAGCATTCTG GCTTACAGGC AGACTCCGTC TTTGATCAGG
                                                                                    600
50
                                                                                    660
       AAAGGAGACC AGTTCAATTT GGCAAAGCTG CACGAAGAGC GTCAGACCAT CAGTGCCCTG
CTGAGAGACA ATGGTTACTA CTACTTCCGC CCACAGGATA TTATCTACGA AGCCGATACC
                                                                                    720
                                                                                    780
      CTCCTCGTAA GAGGTGCCGT ATGCCTGCGA GCCAAGCTCT CGGAAGATAC TCCACCCCAA
GCCATGCGCC CGTGGAGGAT AGGGAAACGG ACAGCAGTCC TGCTCGGAAT GAACGGAGAA
                                                                                    900
55
       AGCCCGACAG ACTCGCTCGA AGTGGAGGAT ATGAAAGTCC TTTACTATCG TAAAATGCCG
                                                                                    960
       GTTCGCCCCA AGATTTTGGC CAAACGCTTT CGTTTCTTCT CCGGCAATCT GTATCGGCAG
                                                                                   1020
       AAAGACGATG AGACGACACG CAAATCCTTG GCTCGTTTGG GAGCCTTCTC CGTTATCGAT
                                                                                   1080
       CTCAATTTT TGCAACGCGA TTCCATTTCC GGCCTTTTGG ATGTGCGACT GCTAACCACC
                                                                                   1140
       CTCGACAAAC CTTGGGATGC ATCATTAGAG ACCTTGTTCA CGAGCAAAAG CAATGACTTC
                                                                                   1200
60
       ATCGGTCCCG GACTGAATTT TGCTCTTGCT CGGCGCAATG TATTCGGCGG AGGAGAAAAT
                                                                                   1260
       CTTTCTTGGA ATATCGGTGG ATCGTATGAG TGGGAGACCG GCAATCGTCC CGAAAATAGC
                                                                                   1320
       AGCAATCGGC TGATCGATAT AAATTCGTAC AACATGAATA CGGCCGTGAA CCTCTCGTTT
                                                                                   1380
       CCCTCGATTG TATTTCCCGG TCTGCTGGAT AAATACTATT ACTACCCCAC GACTACGACT
                                                                                   1440
      TTTCAGGCTT CTGCCACCGC GCTGAACAGG GCACACTACT TTAGCATGTA CTCTTTCGGC
TTTTCGACCA CCTACGAATT TCAGCCCTCC AAGGAACACC GGCATGCTAT TTTCCCGCTC
                                                                                   1500
65
                                                                                   1560
       AAGCTCAACT ACAACCTCCT GGGGCATCAG ACAGAAACTT TCCAGGCCAT TACGGCGAAC
                                                                                   1620
      AATCCGCCCC TGCTGCTCAG CCTTCAGAGT CAGTTCCTTG CTCAAATGGG GTATATCTAT
                                                                                   1680
      ACGTTCAACA AATCCGTTTC AGAGAAAAGT CCTCATCATC TTTGGATGCA ATTCGGACTA TCCGAGGCAG GCAATCTCCT GAATCTGATC TATCTGGCAG CCGGCAAGAA GTACAGCGAC
                                                                                   1740
                                                                                   1800
70
      ACCAAGAATT TCGTCGGCGT CCCCTTCTCT CAGTTCATCA AAGCCACGGG AGAACTGCGC
                                                                                   1860
      TATTCCTATA CCATAGACCG CAATCAGTCA CTGGCAACCC GTTTCGGGAC AGGCGTGATA
                                                                                   1920
      TATAGCTATG GCAATATGCG AGTGGCACCC TATAGCGAGC AGTTCTATGT AGGCGGTGCC
                                                                                   1980
      AATAGTATCA GAGCTTTCAC CGTCCGTAGC ATCGGCCCCG GACGGTTCAA TCCGGATTCC
                                                                                   2040
       GACAATCAGT ATTCCTATTT GGATCAGGTG GGCGAATTCA AACTCGAAGC CAACGTGGAA
                                                                                   2100
75
      TATAGAGGCA AGCTTTTCGG GGATCTCCAC GCAGCCGTTT TCCTCGATGC GGGCAACGTT
                                                                                   2160
```

```
TGGCTCTTGA GGGAGGATTC TTCCCGTCCG GGCGGTGCTC TGTCCGAAGT GGGATCGGTG
                                                                                  2220
       AGCAATTTCC TGAATAGCAT CGCTCTCGGC ACCGGTGTCG GCCTTCGCTA CGATCTGGCA TTTCTCGTGG TTCGTGTCGA TGTCGGCTTC GGTCTCCACC TTCCTTACAA TACGGGTAAG
                                                                                  2280
                                                                                  2340
       AAAGGTTACT ACAATATCCC ACGCTTTAAG GATGCCATCG GTTTCCATTT GGCTGTCGGC
                                                                                  2400
  5
        TATCCCTTC
                                                                                  2409
        (2) INFORMATION FOR SEQ ID NO:63
10
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2349 base pairs
                  (B) TYPE: nucleic acid
                     STRAHDEDNESS: double
                  (D) TOPOLOGY: circular
15
            (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
20
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
25
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...2349
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63
30
       TOTOTOTOCO CGTATATOCG TTTTCCTATG TCCTCGCATT CCGTTCGGTA TCTAATCGGC
                                                                                    60
       ATTGCCGGCT GCTTGCTCCT CATGCTTGCT TCCTCCTGCT CGGTCACCCG TTATGTGCCG
                                                                                  120
       GACGGTAGCA GACTATTAGA CAGGGTAACG ATCGCAAGCG AAACGGGCAG TATCGCTCTG
                                                                                  180
       CCGGAAGATA TTCGGGACTA TACCCTCCAG CAACCCAATT ACAGACTGTT CGGGATGACT
                                                                                   240
35
       CGCTGGCTAC TGCGCGTCTA TAGCAGCTCG AATCCGAACA GCAACAGCTG GTGGAACCGT
                                                                                   300
       TCGCTCCGGA AAATGGGCGA ACCGCCTGTC CTCATCGATT CTGTCCTCAC CGATCGTACT
                                                                                  360
       GCCAACCGTC TGGCAAAGGC GATGGCCGGC GATGGCTTTC TCGATGCTAC TGCTCGTGCC
                                                                                  420
       GTGGTAGACA CCGGCTTGTA CAAGAAAGCT CGCATTACTT ATCTGATTCA GCCCGGAAGC
                                                                                  480
       CGTTATTATA TACGCAATAT GGCTTTGGAT GTGAAGAATC CACTCCTTCC TCCCGTTGCG
                                                                                  540
40
       CTTGGCAATT CGCTTCCTTC GGCATACAAG GTCGGGATCA GCGAGGGTTC TCCCTTGTCG
                                                                                  600
       CCCATCGTAC TCGATGAAGA GAGAAAGGCG ATAGCTCGTC ATATGCGCAA CAACGGCTTC
                                                                                  660
       TGGAAGTTCT CCGCCGAGGA TGTTTATTAT GAAGCAGATA CTACCGTTTC AGGAGGATCG
                                                                                  720
       GGTACGAAAT CTGCCGATCT GAAATTAGTG GTCAATGGCA TCGGGCGTTA TCCATATCGG
                                                                                  780
       ATCGGCAGGG TATTCTTTCA TGCCGATTAT GATCCTCTCG AATCGGACTT CAGAGTTCAG
                                                                                  840
45
       GAGCTGCCAC GTATCGATTC GATTTCGCGT GGCGATTACA CTGTTTACTA TGGGAGTAGG
GGACGTTATA TCCGGGCATC GGCTCTCACG CGGTCGGTGT CCGTTACACC GGGAGCTTTT
                                                                                  900
                                                                                  960
       TTCTGCGAGG ATGATGTGGA ACGCTCTTAT ATCAAGCTGA ATGCGCTCCC TATCGTTCGG
                                                                                 1020
       AACGTGAATA TCCGATTTGT GGAGCACAAT GGTAAGGATG AGATTGCTCT GGCGGATAGC
                                                                                 1080
       TCTCGCCTTG TGGACTGCTA TATTCTTACC GTTCCGGCCA AGAGCAAATC GTTCGAAGCC
                                                                                 1140
50
       GAAGTCCTCG GCACCAATTC CGCTGGAGAC TTCGGGGCGG CTTTGTCTCT CGGTTTCACC
                                                                                 1200
       GATCGCAATT TGTTTCGTGG GGCGGAGATG TTCAATATCA AACTCAAGGG TGCTTACGAA
                                                                                 1260
       GCCATTCGCA AGGGTTCGCA CAGCTTCATG GAATATGGGG TGGAAAGCTC GCTCCGTTTC
                                                                                 1329
       CCTCGTCTCC TCTTCCCATT CATTTCTGAC GAAACGCGCC GGCGGCTACG GGCATCCACG
                                                                                 1380
      GAATGGAAGA TCGGGTATAA TTACCAGACA CGTCCGGAGT TTGATCGGGT GATTCTCTCC
GCTCAACTCA ATTATTCATG GCAGACCTAC CTGCACAATC GTCTGCGTCA TACGATCCGC
                                                                                 1440
55
                                                                                 1500
       CTGCTGGATG TCGATTATCT CCATCTCCCG TACATCGATC CCGACTTCGC CCAATCCCTT
                                                                                 1560
       CCGCCTACGA CTGCACTGTA TAACTACACG GAGCAGTTTA TCCTCGGCTC GGCATATATA
                                                                                 1620
       CTGAACTATA CCACGGCTTC GTCCATGGAG CGTACCGTAT CCAATCCTTT TACGGCACGG
                                                                                 1680
      TTCAGTATCC AGACAGCCGG CAACCTGCTG CAAGCCATTT CTTATCTGAC CGATTCTCCG
                                                                                 1740
60
      AAAGACGAAC ACGGGTTGTA TAAAATGTTC GGTCTGCACT ATGCTCAGTT CGTCAAGCTC
                                                                                 1800
      1860
                                                                                 1920
      TTTGCCGGAG GATCGAACAG CGTTCGCGGC TGGAGTGTCC GTACCCTCGG CCCGGGGAGT
                                                                                 1980
      ATGAAGATGA CTCCGGACAA GACCTTCTTC GATCAGATGG GTGATATTCG TCTGGATCTG
                                                                                 2040
65
      AATGTCGAAT ACAGGACAAA GCTGTTCTGG AAGTTTCGCG CAGCAGCTTT TGTCGATGCC
                                                                                 2100
      GGCAATGTCT GGACGATAAA GGAGTATGAG AATCAGGAGG ACGGTCTCTT TCGTTTCGAT
                                                                                 2160
      CGCTTCTACA AGGAAATAGC TTTGGCCTAC GGTCTGGGGC TTCGTCTGA CTTCGATTAT
TTCCTTGTGC GGCTGGATGC CGGACTGAAA GCCTACGATC CTCAGCAGAC AGGGCGTTAC
                                                                                 2220
                                                                                 2280
      AAATGGGCTA TCACACGCCC AAACCTTTCT TCCAATTTCG CTTGGCACAT TGCAGTAGGC
                                                                                 2340
70
      TATCCGTTC
                                                                                 2349
```

(2) INFORMATION FOR SEQ ID NO:64

75 (i) SEQUENCE CHARACTERISTICS:

```
(A) LENGTH: 2625 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
 5
          (ii) HOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
10
          (iv) ANTI-SEUSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
15
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...2625
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64
20
      GTCGAATCTA AATTGTTATG TCTTATGAGA AAAAGAATTC TACAACTTTT CCTGACCGCA
                                                                              60
      TTGCTGCTGG CATTAGGCTC CTCTCTCGCC ATAGCGCAAA CAGTGGTGAC CGGTAAGGTG
                                                                             120
      ATCGATTCAG AAACGTCCGA ACCGCTCATC GGTGTATCCG TAAGCACCGG TCAGGGAGCA
                                                                             180
      TCCCTCCGCG GTGTAACCAC CGATATGGAT GGTGGCTTCC GATTCGAAGT ACCGGCCAAA
                                                                             240
25
      TCTGTCTTGA CTTTCCGTTG CGTAGGTTAT GCTACCGTAA CTCGCTCTAT AGGCAGAGGT
      TCTCAAGAAG ACCTCGGTAC GATCTCCTC GATCCCCAGG CCATCGGCTT GGATGAGATT
                                                                             360
      CAGGTAATAG CCTCTGTGGT GCCCAAAGAC CGTATGACGC CGGTACCCGT TTCCAATATC
                                                                             420
      CGTGTGGCTG ATATTCAGGC AGCATCGTTG AATGTCGAAT TTCCCGAACT GGTTAAATCC
                                                                             480
      ACTCCCTCTA CCTATACGAC AAAAGGAAGC GGAGGTTTCG GTGATGGTCG TACCAATGTG
                                                                             540
30
      CGTGGATTCG ACACTTACAA CTTCGGTGTA CTCATCAACG GAGTTCCTGT CAATGGTATG
                                                                             600
      GAAGACGGGA AAGTATATTG GAGCAATTGG AGTGGTCTGA TGAATCAAGC CAGTACCATT
                                                                             660
      CAGATTCAGC GCGGACTCGG AGCCTCCAAG CTCGGTATCA GCTCGGTAGG TGGTACGATG
                                                                             720
      AACATTATCA CGAAGACTAC GGACGCCAAC ACCGGAGGTT CGGCTTATGT CGGTATGGGT
                                                                             780
      AATGATGGAT TGCACAAAGA ATCGTTCTCC ATTTCTACGG GTATGAACGA CGGTTGGGCT
                                                                             840
35
      ATCACCATTG CAGGCTCCCA TATGACGGGT CTGGGTTATG TGAAGGGGCT GAAGGGACGT
                                                                             900
      GCATTCTCTT ACTTCTTCAA CGTTTCGAAG AAGTTCAATG AACGTCATAC CCTCTCTT
                                                                             960
      ACCGGATTCG GTGCACCACA ATGGCACAAC CAACGTTCTT CCAAATATTC TGTAGCCGAC
                                                                            1020
      TATGACAAAT ACGGCATCCG TCACAATCAA TCCTTCGGCT ATCTGCGAGG CGAACTGACT
                                                                            1080
      CCTACGGCTT ATGCTTACAA TACGTACCAC AAGCCCCAGT TCTCGCTGAA CCACTTCTGG
                                                                            1140
40
      AAGATGGATG AAAATACCTC TCTTTATACc gCAHCCTACG CATCTTTGGC TACCGGTGGA
                                                                            1200
      GGTCGTCGCG CTTATGGAAA GAACAGTAAG TGGGTATTGA TCAACTACAA CACCGGACAA
                                                                            1260
      CCCTATGAAC AAACAAAGGT GACTCCCGAT GGACTTATCG ACTACGATGC CGTACTGGCT
                                                                            1320
      GCCAATGCTG CGGCGAGCAA TGGCTCGGAA GCAATTTTTG CCCTTGGCTC CAACTCTCAC
                                                                            1380
      AAGTGGTTCG GTCTACTCTC TTCATTCAAG AAGAAACTTA ATAGTTCGCT GACTTTGACA
                                                                            1440
45
      GCCGGATACG ATGGGCGTTA CTACCGTGGC GACCACTATG ACAAGATCAC CGATCTGCTC
                                                                            1500
      GGCGGTAGCT ACTACATAGA GGATCCCAAG ACAAAGCTCG CATACCATGC GGAAGGTCAG
                                                                            1560
      CAACTGAAAG TGGGTGACAT TGTAAATCGG GACTACACAG GCGAAATCAT GTGGCACGGC
                                                                            1620
      CTCTTCGCAC AGATGGAGCA TTCGTCCGAA TGGATCGATG CATTCGTATC AGGATCTATC
                                                                            1680
      AACTACGAAC TATACCGCAA TCACAACTAT GGCGGTAGCA AGTCCACCGG CTACCTGCCC
                                                                            1740
50
      GGCGTATCGC CGTGGAAAAG CTTCCTTCCG TGGAGTGGCA AGGCAGGTCT GAGCTACAAG
                                                                            1800
      TTCGCACAGG GACACAATGT ATTCGCCAAT GGCGGTTTCT TCACACGTGC ACCACTCTT
                                                                            1860
      GGCAATATCT ATGCTGCGGG GGCTATCATT CCCAATGACA AAGCCAATAT GGAAAAGGTG
                                                                            1920
      CTTACAGGAG AGGTCGGCTA TGGATTCACG AATCACAAAA ACTTCGAGTT CAATATCAAC
                                                                            1980
      GGATACTATA CGAAGTGGAT GGATCGCGTG ACCTCGAAGA GAATCGGAAA CGAGTATGTT
                                                                            2040
55
      TATCTCAATG GCGTTGATGC TGTTCACTGT GGGGTAGAGG CTGAGGTCAG CTATCGTCCT
                                                                            2100
      ATTCGTCAGA TCGACCTTCG CGGTATGTTC TCTCTCGGTG ACTGGACTTG GCAAAACAAT
                                                                            2160
      GTAAGTTACA CTTCTTACGA CGAAGCCGGC AATGAGACAG GGCAGGATAT AACCTATATC
                                                                            2220
      AAGGGTCTTC ACGTCGGAGA TGCAGCACAG ATGACGGCTG CTGTATCGGC AGACATAGAG
                                                                            2280
      CTGTTCAAGG GTTTCCATGT CATAGGTAAG TACAACTTCC TTGGCAAGAA CTATGCAGGA
                                                                            2340
60
      TTCAACCCCG CAACGCGTAA TGCACAGCAG TACGAAGCGG ATGGCAAAGA AATCGTGGAA
                                                                            2400
      TCATGGAAGT TGCCCGATGT AGGTCTGTTC GATCTGTCTG CATCCTACAA TTTCAAGCTT
                                                                            2460
      GGTTCACTCA GCACCACATT CTATTTCAAC ATGGACAACG TAGCCGACAA GCGATATGTG
                                                                            2520
      AGCGATGCCG ACGACAATAT CATCGGTAAG AAACACGATG AGGCTTCGGC TCTCGTATGG
                                                                            2580
      TACGGTTTCG GCCGCACTTG GTCTACCGGT ATTCGTGTAA ACTTC
                                                                            2625
65
      (2) INFORMATION FOR SEQ ID NO:65
           (i) SEQUENCE CHARACTERISTICS:
70
                (A) LENGTH: 1380 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
75
```

(ii) MOLECULE TYPE: DNA (genomic)

WO 99/29870 PCT/AU98/01023

44 / 490

```
(iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
  5
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
 10
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1380
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65
 15
       AACAGATGTA ATCCCTCGCG TCAATACTTC CACCTATCGC AGAATGACGT TGGAAGATCT
                                                                                   60
       CTCCGAGCTT TTCTAATCCA TCTATCTATG AAGTTTTCAA TCCGCCTTTT CCTCTGCATC
                                                                                  120
       ATCTTTCTCC TCTCTGCATT TATCCTGCCT GCTCTCGGAC AAAAATCCAA GCAGGTACAG
                                                                                  180
       CGACTTGAGA AGCAACGTAA GGAGGCCCTC AAAGCCATCG AAAAAACCGA TCGCGAACTA
                                                                                  240
       CGAAATACCA AGAAAGACAA GCAAGACAAA CAAAAGCATC TCAACCTCCT GAACAAGCAG
                                                                                  300
20
       GTTGCTCAAC GCAAGCAGAT GGTACAACTC TTGGACAATG AGGTCAAAGA GTTGCAATCC
                                                                                  360
       GACATTGATT CCATGACGGG TGTATGTCAT CAGCTCTCTG TAGAAGAGAA AGCCCGATCC
                                                                                  420
       GATGAATATG CCCAAGCTCT ACAGTCTATG CAAAAGCGGA AACGCTCGTT GGATCGCATC
                                                                                  480
       CTTTTCATTT CATCGGCCAA GAGCTTTGAC GAAGGCATGC GACGGATGCG TTTCTTGGAA
CAATACGCTT CTGCATACAA GCTGGCATCT GTCCGGCTGC GCGATACACG TAGCAAGTTG
                                                                                  540
                                                                                  600
25
       GAGACTGAAC GTGCGACTGT AGAAGACGCC AAAAAGGAGA AAGGACATCT CTTAGTCATC
                                                                                  660
       AGAGAAGAGG AAAAAAAGAA ACTCGAAGGA CAGCAAGCCG AGCAACGTCG GCAGGTGCAG
                                                                                  720
       GCTTTGGGAG CCAAACAAAA AGACTTGGAA GCGCAGCTGC GAAAGCAGAA AAAGCAAGCC
                                                                                  780
       GAAGCTCTGA ACAGAAAGAT CGAGAAACAG ATTGCCAAGG AAATAGAAGC TGCCGAACGT
                                                                                  840
       CGTGCTCGAG AAGAACGTGA ACGGTTGGCA CGCGAAGCCA AAGCCAAGGG TAAGCCGGTT
                                                                                  900
30
       CCTGCCGAAC CGGAACGGAA GGCGGAGACC AAAGGCGGCT ATGCTATGGA TGCCTCTGAG
                                                                                  960
       CGTGCTCTCT CGGGCAGCTT TGCACAGAAC AAAGGTCGCC TGCCCGGCCC CGTTCGCGGC
                                                                                 1020
       AGATACCGAA TCGTAAGCGA CTTTGGCGTG CATCAGCACA GTGAGCTGAA AAAAGTACAA
                                                                                 1080
       GTTAATAATG GAGGTATCGA CATCGCTGTA GCAACAGGAT CCGATGCTAC CAGCGTATTC
                                                                                 1140
       GATGGTGTAG TGTCCAGTGT ATTCGTGATA CCCGGTTATA ATTCGGCCGT AATGGTTCGT
                                                                                 1200
35
       CACGGTAACT ATATCACGGT TTATGCGAAT CTGAGCAAAG TGTATGTAAA TTCCGGCACT
                                                                                1260
       CGTGTTAAAA CGGGTCAGGC TCTTGGTCGT GCCTATACGG ATCCTTCCAA CAACCAGACC
                                                                                1320
       ATTATTCACT TCGAAATCTG GAAAGAACGC AGCAAACAAA ACCCAAGACT ATGGTTACGA
                                                                                1380
40
       (2) INFORMATION FOR SEQ ID NO:66
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1026 base pairs
                  (B) TYPE: nucleic acid
45
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
50
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
55
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1026
60
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66
      AGTTTTTATC AAGAAATAGA CAGACTTATG AAAAAGTATT TGTTATATGC CTCGTTGCTA
      ACGAGTGTTT TGCTCTTTTC CTGTTCAAAG AACAATCCTA ACGAGCCGGT GGAAGACAGA
65
      TCCATCGAAA TTTCTATAAG GGTAGATGAT TTCACCAAAA CGGGTGAGGC AGTACGCTAT
                                                                                 180
      GAAAGGAATC AAGGAAGTGC TGCCGAAAGG CTCATTACCA ATCTTTACCT CTTGTTGTTC
                                                                                 240
      GATCAGTCAG GGGCGAATCC GGCGAAATAC TATATTACCG GTAACACTTT CACCGGAGGG
                                                                                 300
      ACCTGGCTTC CTGACGATAT GAAGGTGAAG TTGGATATGA CACAATCCGA GGCCGGAGAG
CGCAAAGTAT ATGTCGTAGC CAATGTTGAT AATGCGGTTA AAACGGCTCT TGATGCTGTC
                                                                                 360
                                                                                 420
70
      GCTAACGAAA GCGATTTGCA GACTGTAAAG AGGACGACTG CAATGCCGTG GTCGACCGAT
                                                                                 480
      ATAGCCTCTC CTTTCCTGAT GTCCGGAAAC AAGACACACG ACTTCTTGGC CAATCGTCTT
                                                                                 540
      TTGGACAATG TGCCCCTTGT GCGTGCCATT GCCAAGGTGG AGCTGAATAT CTCGCTGAGT
```

GAGAAATTTC AGATTGTGCC GATAATTGTC AATGGTAGTT TGAGTGAGTT CAAGTTCAGA

TACGTAAACT TCGACAAGGA GACCTACGTA GTGAAGCCAA CGACCAAGCC GGACAATCTC

ATTAGTTCTG CTAATGGTGT TTGGCCTCAG ATTACAGATT GGACTGTATG GGGTGCTTCC

600

660

720

780

75

| 5 | TTAAATACTT CTCCTGCTCC GGATGCGGGC ACAGGTTATA CATTGGATGC AAATGGCAAG GTAACGGCAC TACGGATGT TACCTATCTG AATGAGCGCC ATAGCAAAGG GGCTACGGTA GAGGTCGCAT TGCCTCGTGT GGATGATGGC ACCCTTCCTC CTCCGGAATT CGGTCCGGAG CTTTATCGTT TGCCTTTGCC GGACAAGATC CTGCGCAATC ATTGGTACAA GTATGAAGTC GAGATT | 840 900 960 1020 1026 |
|----|--|-----------------------------------|
| | (2) INFORMATION FOR SEQ ID NO:67 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 987 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 15 | (ii) NOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 20 | (iv) ANTI-SEMSE: NO | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROHONAS GINGIVALIS | |
| 25 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1987 | |
| 30 | (mi) SEQUENCE DESCRIPTION: SEQ ID NO:67 | |
| 00 | AACCTTAGGA CACAGCCTTT CTTTTTGGTA GATTTGCAAT CTATGATCAG AACGATACTT TCACGATATG TATCCTCGAA CTTTTGGAGT CGGGGAGCTA CCTTTTTTTT CACGATTTTC CCGGCCTTCA TCCTCGCCGC TACTGCTTTG CCGGCTTGTG GAGGGGGTAC TGCTTCAGGC | 60 120 180 |
| 35 | TCCGATCGTA CGCTGGCTGT GACCATCGAG CCACAGAAAT ACTTCATCGA GTCCATTGCG GATAAGTCGG TGCAGTGGT GGCATTGGTA CCGGCCGGCA GCAATCCGGA GGAATACGAC CCTTCCCCTA CCCTGCTGCA CCACAGCAGCA GCAATCCGCA GCAATCCCAATCCAAT | 240 300 |
| | CCTTCGCCTA CCGTGATGAA GCGTTTGTCC GAAGCAGATG CCTACTTCTA TATAGGAGGA CTGGGGTTCG AGCAAAGAAA TCTCGCTGCC ATTCGGGACA ATAACCCTAA GCTCCCTCTT TTCGAAATGG GCAAAGCCTT GGCGGATGCC GGAAGTGCAG ATCTCCACGG CTCCTGCACA | 360 420 480 |
| 40 | GATCATTCTC ATACAGACCT GCATGCCCAT GATCCGCACT ATTGGAGCAG TGTGGTAGGG GCAAAGGCAC TCAGTCGTGC TGCATACGAC GCGCTTGTGG AGCTTTATCC GAACGAGAAA GACAAATGGG ACAAAGGGCA CGACCGTCTC AACGGACGTA TCGACAGCGT GAAGAGACTC GTCGATACCA TGTTTGCCAA TGGCAAAGCA GACAAAGCCT TCGTCATATA TCACCCATCG | 540 600 660 720 |
| 45 | CTCAGCTTTT TCGCCCAAGA GTTCGGCCTG CGGCAGATCG TCATAGAGGA AGATGGGAAA GAGCCTACGG CTGCCCACCT TCGTCGTGTG ATCGATCAGG CACGTGCCGA TGGTGTCAGA ATCGTATTTA TCCAACCCGA ATTTGAAACG CGTCAGGCGG AGGACATCGC ACGCGAGATC GGTGCTCGTC CGGTAAGGAT CAATCCTCTG CGCAGCTCGT GGGAGGAGGA AATTTTACAT ATTGCTCGCG CTTTGGCTCA TGAACGG | 780 840 900 960 987 |
| 50 | (2) INFORMATION FOR SEQ ID NO:68 | |
| 55 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2634 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 60 | (iii) HYPOTHETICAL: MO | |
| | (iv) AUTI-SEUSE: NO | |
| 65 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYRONOMAS GINGIVALIS | |
| 70 | (ix) FEATURE: (A) NAHE/KEY: misc_feature (B) LOCATION 12634 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68 | |
| 75 | GCAGATTCTA TTCGATATCC TCTTTACTTT TTTGGGCGGA ATCGGAAGAA ATGCTTTAGG GAACCTATTC CCACCTTATA CAATAAAAAC ATGATCGGAA AAAAAATCTT TTTTATCCTG CTGGCGCTCA TTGCGTTCAG TGGGCTGAAC GCAGCGACAG ACACTGAGTT CAAGTACCCG | 60 120 180 |

```
ACCGATGCCA ATATCATCGG TCACGTCAAA GACAGCAAGA CGGGTGAACA CCTTGTCGGT
       ATCACTATTG CTATCAAAGG CACTACCTTT GGTACATCTA CAGATGCAAC CGGGCACTAC
       TATCTTCGTA ACTTGCGTCC GGGTGAGATC ACTTTGATTA TGCGTGGCAT GGGCTATAAG
                                                                                 360
      AGCCAGGAGC GCGTAGTCCG CGTAGAAAAG GACAAGACTA TCGAGGTGAA TTTCGAAGCA
                                                                                 420
 5
       GAAGAGGATG CCATCAATCT GGACGAAGTC GTGATTTCGG CCAACCGCGA ACTGACGCTT
                                                                                 480
       CGCCGTCTTG CTCCTACTCT GGTAAATGTA TTGAACGAAA AAGTCTTCTC GCAAGTCAAT
                                                                                 540
       GCTTCTAACC TGGCTCAAGG CTTGTCATTC CAGCCGGGAG TTCGTGTAGA GAACAACTGT
                                                                                 600
      CAGAACTGTG GTTTCAATCA AGTTCGTATC AATGGACTGG ATGGTCGTTA TGCACAGATC
                                                                                 660
      CTCATCGACA GCCGTCCCAT CATGAGTGCC CTTGCCGGTG TTTACGGTCT GGAGCAGATC
                                                                                 720
      CCTGCCAATA TGATCGAACG TGTGGAGGTA GTACGTGGTG GAGGATCGGC CTTGTACGGT
10
                                                                                 780
      TCTTCTGCTA TTGCCGGAGT GGTGAATATC ATCACCAAGG AACCTTCTCA CAATTCTTTC
                                                                                840
      ACATTCAATG AATCTCTGAG CTTTACCGGT TTCAGCAAGC TGGATAACAA CACGAACTTC
                                                                                900
      AATGCCTCCA TCGTCAGCGA TGACAACCGT GCCGGTGCCA TGGTATTCGG GCAGGCTCGT TACCGCAACC ATTGGGATGC TAACAATGAC GGTTATTCCG AATTGGGTAA AATAGATGCC CGCTCGCTGG GAGCGCATTC TTATTTGCGC TTGAGCGACT ACAGCAAATT GACGGGAGAG
                                                                                960
                                                                                1020
15
                                                                               1080
       TTTCACACGA TCAGTGAATT CCGCCGTGGT GGCGATCGTA TCGATTTGCC TCCTCACGTA
                                                                               1140
      GTGGGTGTAG CTGAACAAC TGACCATAGC GTATTTAGCG GAAACTTGAA ATACGATCTC
                                                                               1200
       TTCTCTTCCA ACTATAAACA CCACTTCCAG GCTTATACTT CCGGACAGAT CGTAAATCGC
                                                                               1260
      AAGAGCTATT ACGGAGGTAT CGGAGAGATT GACGTCAATG GCCACCCCGG TGGTACGGAA
                                                                               1320
20
      GGCTACCCTA TCCCTCAAGA TCAATACGGC AATAATTATG GCGTGACCAA AGGCAAGACA
      TATATGGGCG GTATCCAGTA CAGCTACGAC TTGGACAAAT TCCTCCTCAT GCCTTCGCAA
                                                                               1440
       CTTTTGTTCG GAGCCGAATA TACGCGTGAT GAACTCAATG AMGTGATGCC CATCCTTTCA
                                                                               1500
       TGGCAGACCG GCGAGGATGC CAATGGGAAT ACCATTCCCC TCTATCCCGA ATTGGATCAG
                                                                               1560
      AATATCAACA ACTACAGCCT ATTCGGTCAG AACGAATGGA AAAATGACAG ATGGAGCATC
                                                                               1620
25
       CTTGTTGGCG CTCGCTTGGA CAAGCATAGC GAAGTCAAGG ATATGATTCT GAGTCCTCGT
                                                                               1680
      ACCACACTGC GTTTCAACGT GAATCCGGAC ATCAACCTGC GCGCTACATA TGCAAAAGGG
                                                                               1740
       TTCCGCGCAC CGCAGGTATT CGATGAAGAC TTGCACGTAG GGGTTGTAGG CGGTGAGGCA
                                                                               1800
      CAGAAAGTAT TCAACGATCC GAACCTCAAG CCTGAAATTT CTCATGCATT CAGTTTGAGT
                                                                               1860
      GCCGATATGT ATCATCGTTT CGGTAACGTC CAGACCAACT TCCTTGTGGA AGGCTTCTAT
                                                                               1920
30
      ACTCGTTTGC TGGATGTATT CACCAACGAG GAGCAGCCTG ATCAGCACGA TGGCATCAAA
                                                                               1980
      CGCTACACGC GTATCAACGG TAGCGGAGCC AAAGTATTCG GTCTCAATCT GGAAGGTAAG
                                                                               2040
      GTCGCATACA AGTCCTTCCA GCTCCAAGCC GGTCTTACCC TGGCCAGCAA CAAATACGAC
                                                                               2100
      GAAGCACAGG AGTGGGGTCT GAATACGGTG AAAGACACCA ACGGAGCTTT TGTTACCGAG
                                                                               2160
      GCCAATGCAA ATGGACAACA GGAATACAAG AACGAATCCA TGACGGATAC GCAGATCACC
                                                                               2220
35
      CGTACCCCCA GCGTATACGG TTATTTTACT TTGGCCTACA ATCCTGCTCA CTCATGGAAC
                                                                               2280
      ATAGCCCTTA CGGGAGCATA TACCGGTCAG ATGTATGTAC CCCACGCTAT CGAATATGGT
                                                                                2340
      GTGAAGTCTG CCGAACTGGA TATTATGCAG AACAATCCTG AGATTACCGA CGAAACCGGA
                                                                               2400
      AAGGCTCCCC GTATTGATGA GCTGAAGAAG ACACCTGCAT TCTTCGATTT GGGCTTGAAA
      GTGGGTTATG ACTTCCACGT ATTCCAGGCT ACTGAGGTTC AACTCTATGT AGGTATGAAC
                                                                                2520
40
      AATATCTTCA ACTCTTTCCA GAAGGACTTC GATCGTGGAG CTGCACGTGA CAGCGGATAT
                                                                                2580
      ATCTATGGTC CTACGCAGCC GCGTACAGGC TACATGGGCT TGGTAGTGAA GTTC
                                                                               2634
      (2) INFORMATION FOR SEQ ID NO:69
45
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 618 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
50
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
55
          (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
60
           (ix) FEATURE:
                 (A) NAME/KET: misc feature
                 (B) LOCATION 1...618
65
           (Mi) SEQUENCE DESCRIPTION: SEQ ID NO:69
      AAACAGATAG TTATGACAGT AAAGCGCGCA GTGCGAATAG CACTTCTCAC GCTGATAGGC
      ATTCTTTTTT CCTCACCTTC TCTTGTTCGG GCGCAAAGTC TTTTCAGCAC CGAACATGTC
                                                                                120
      TTGCAACTAT ACAACAAGAT ACTCTATGGA GAGTCGGCGG CGGATACCGT CGCAGAGAAA
                                                                                180
70
      ACGGCAGGTG AGTCGGCATT TCCTTTTATA GACAAACTCA TCAATCTCGG CCGCACTTTC
                                                                                240
      CTCGGCAAAC CATATCGCTA TCGCGGTCCT TCCCCATGGC CGATGGACTG CTCGGGCTAT
                                                                                300
      GTGTCTTACC TCTACTCCAA ATTCGACATC AAACTCCCAC GTGGTGCGGC AGCACAGAGC
                                                                                360
      CAATATACGA ATCCTATCGA GCGCGAGGAT GTTCGTCCGG GCGACCTCCT TTTTTTCAAA
                                                                                420
      GGCCGCAATG CACGCAGCAA CCGTATCGGG CATGTAGCTT TGGTCGTATC TGTCGATGAA
                                                                                480
75
      GATGATATTA CCATGATGCA CAGCCGCAAT TCGCGAGGGA TCGTGATCGA AAAACTCAAT
                                                                                540
```

| | CGCAGTGCAT ACTTCTCCCG TCGCTTGGTG AGCTATGGCA GGGTACCCGG AGCCAAGAGA GTGATCCCAC GAAAAAGT | 600 618 |
|----|--|--|
| 5 | (2) INFORMATION FOR SEQ ID NO:70 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1401 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 15 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 25 | <pre>(ix) FEATURE: (A) NAHE/KEY: misc_feature (B) LOCATION 11401</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70 | |
| 30 | AAAGGTACGT GGAATAGAAA AAACCGAAGA GAAGAAATGA AACGGACAAT CCTCCTGACG GCACTGACG TCCTATCTC GCTCTCCTTG CTTCGTGCAC AAAATGAATC CGAAGCATCA ACCAATCCGA TGTCAGGCCT CTCCCTGGAA GACTGTATCC GGATAGCCAA GGAGCGCAAC CTGAATCTG GCAGCAGGA GATCGAACAA GAAAACCGAA TCATTAGTCT CGATGCAGCA CGACACAGTT TCCTGCCCTC GGTCAATGCA GGCATCGGAC ACAACTATAG CTTCGGACGT TCGAAAGACA AAACGGGAGT ACCGTAGAT CGCTCCTCCAG TGAATACCAA TCTCAGCATC | 60 120 180 240 300 360 |
| 35 | GGAGCTTCGG TGGAAGTATT CAGCGGCACA CGTCGTCTGC ACGACCTCAA GCAGCAAAAG TACAACGTGG AGGATGGTAT AGCCCGACTT CAAAAAGCGC GTGAAGACCT CAGCCTGCAA ATCGCGGCTC TCTATATCAA TTTGCTCTTC CGTCAGGAAA TGACTCGTAC GGCAGAAACA CAGTTGGCAC TGATTCGCGA GCAACGCAAT CGCACGGCCG AAATGGTTCG CGTAGGTAAA TGGGCAGAGG GTAAGCTCCT CGACATAAAT GCCCAGGATGG CCAAGGACGA ACAACTTCTC | 420 480 540 600 660 |
| 40 | GTACAATATC GTTCGGAGGA GGAGCTGGCT CGTCTGGACT TGGGGCAAGC CCTCGAACTG GAGCACCCCG AAAGCATTGC AGTCAAGGCT CCCGACACAG ACGTTCTCGT AGCAGAAAGG TTGGGATCTC TCCTTGCTCC CGAAGAGATC TATCGCACGG CTCTCGGCTT GAAACCGGCA CTGCATCCGA GCGAGCTGCA AATAGCTTCG GCACGCGAAG GTCTGGCCTC GGCTCGTGCG GCATACTTCC CGACGCCCAG CCTCTCTGCC GGATACAGCA ACGGTTACTT CCGCGACCTC CCGACGCCACAC ATGCCCCAM ATGCCCCAM ATGCCCACAC ACGCTACTTC CGCGACCTC | 720 780 840 900 960 |
| 45 | GGCAAGGAGT ATGCCGCCAT CAACCCCTCC TTCTCCGAAC AGTGGAAGAA CAACGGCAGC TACAGTATCG GACTCTCTTT GAATATCCCC ATCTTCTCTG CCATGCAAAC GCAAGATCGC GTTCGGAGCA GTCGCCTGCA AAAACGCTCA AGCGAGCTTC GACTCGTCGA AGAGAAAAAA GCCTCTATA AAGAGATCAG GCAAGCATAC AGCAATGCCG TGGCAGCCGA TAAGGCCATC GCAGCAGAGCAG | 1020 1080 1140 1200 1260 1320 |
| 50 | AGCCAAGTGG AAGAACTTCG TGCCAAGTAT GACTTCATAT ACAAAGCCAA AGTTTTGGAT TTCTATCAGG GCAAAGACTT C | 1380 1401 |
| 55 | (2) INFORMATION FOR SEQ ID NO:71 (i) SEQUENCE CHARACTERISTICS: | |
| 60 | (A) LENGTH: 1353 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 65 | (iii) HTPOTHETICAL: NO | |
| | (iv) AUTI-SENSE: NO | |
| 70 | <pre>(vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE:</pre> | |
| | (A) NAME/KEY: misc_feature (B) LOCATION 11353 | |
| 75 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71 | |

```
AAAACATCAT ATCGGAATAA TATGCGTTTC CAACATTATC TCATCTGTAC GGCTGCCGTA
GCGGCTTTGG CTGCGAATCC CCTTACGGGC CAATCGAATA TGACCCTCGA AGAGTGCATA
                                                                                     60
                                                                                    120
       GACTATGCAC GCCGGCACAG TTCGGCCGTG GCGCTGTCCG CTGCGGAACT GGAGCAGTCC
                                                                                    180
  5
       AAGGCCGATT ACCTTCAGGC CGTCGGCAAT TTTCTGCCCC GTGTATCGGC CGGAACCGGT
                                                                                    240
       GCTTCGTGGA ATTTCGGACG CGGATTGGAT GCCGAGACGA ATACCTACAC CGACATCAAC
                                                                                    300
       AGCTTCAACA ATTCGTACAG CATACATGCC ACGATGACCC TTTTCGACGG TTTGCAGAGT
                                                                                    360
       GTCTATCGAC TGCGGATGGC GCATGCACGC CGGGAGGCTT CGCGCCTCTC CGTTCGCGAG
                                                                                    420
       CAGCAGGAGC TGGCAGCTCT CGGCACCACG GAGGCCTACT ACGACCTCGT CTATGCGCGC
                                                                                    480
10
       CAAATGCAAG AGCTGGCCAT GCAGAAGTAC GAGGAGAGCA GCCGCCTCCA CCGGCAGACG
                                                                                    540
       GCTCGAATGG AAGAGCTGGG GATGAAGAGT CGTCCCGATG TCCTCGAGAT GCAGTCGCGA
                                                                                    600
       ATGGCCGGTG ACCGTTTGGC CCTGACTCAA GCGGACAATC AGTGCATCAT CGCTCTGATC
                                                                                    660
       CGGCTCAAAG AAAAAATGAA CTTCCCCATC GATGACGAAC TCGTCGTAGA CGATATGCCG
                                                                                    720
       GCTGACAGTC TCTCCGCCGA CATGGCCGAA TCGGACAGCT CGGCCGGCGT CTTCGCCCGT
                                                                                    780
15
       GCTGCCCATC ATCATCCCGT CCTCCTCCGT GCCAAACTCG ACGAGCAGGC TGCCACCGAC
                                                                                    840
       CGTTTGCGAG CCGCGCGAGG TGCATTCCTG CCGAGTGTGT CGGTATCCGG AGGATGGAAC
                                                                                    900
       ACGGGATTCT CACGCTTTTT GAATGGATCG GACTATACGC CCTTCAGCGA GCAGTTTCGG
                                                                                    960
       AACCGTCGGG GGGAATACGT CAGTCTGAAT CTGAGTATCC CCATCTTTTC GGGATTCAGC
                                                                                   1020
       CTTGTGAGCC ATCTGCGTCA GGCGCGTGCC GAACGCAGGG CGGCAATCGT CCGACGGGGC
                                                                                   1080
20
       GAAGCGGAGC GCAGGCTCTA CAGCGAGATC GCCCAAGCCA TGGCCGACCG GGATGCCGCT
                                                                                   1140
       CTGGCTTCCT ACCGCCAGGC GAAGGAGCAT ACCGACGCCA TGCAAACCGC TTACGAAGCC
                                                                                   1200
       GTCTTGCAGC GTTATGAGGA GGGGCTGAAT ACGGCCATCG ACCTGACCAC TCAGGCCAAT
                                                                                  1260
       CGGCTCCTGG ATGCCCGTGT GCAGCGACTG AGAGCGGCCA TGACCTACCG GCTCAAATGC
                                                                                  1320
       AAACTCATAG CCTATTACGG CTGCCTTTCG GAC
                                                                                  1353
25
       (2) INFORMATION FOR SEQ ID NO:72
             (i) SEQUENCE CHARACTERISTICS:
30
                  (A) LENGTH: 2886 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
35
           (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
40
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
45
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2886
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72
50
       GCCATTTTTG TCGTATCATT GCAAATTGAA AAAATAACAG AGAATAAGTA TAATTCAGAC
       AAGAGCATGA ACAAATTTTA CAAATCACTT TTGCAGTCAG GACTGGCTGC CTTCGTGTCG
                                                                                   120
       ATGGCAACTG CACTGACCGC TTCTGCACAG ATTTCGTTCG GAGGGGAACC CTTGAGTTTC
                                                                                   180
       TCTTCAAGAT CCGCCGGAAC GCATTCATTC GACGATGCAA TGACTATCCG CCTTACTCCG
                                                                                   240
       GATTTCAATC CGGAAGACCT GATCGCACAG AGCCGTTGGC AATCGCAAAG AGATGGCCGG
                                                                                   300
55
       CCCGTCCGGA TAGGACAAGT AATACCGGTG GATGTGGACT TTGCATCCAA GGCTTCGCAC
                                                                                   360
       ATCTCTTCCA TCGGAGACGT AGATGTATAT CGCCTGCAAT TCAAGTTGGA AGGAGCCAAA
                                                                                   420
       GCCATTACGC TTTATTACGA TGCATTCAAT ATTCCGGAGG GCGGACGCCT CTATATCTAT
                                                                                   480
       ACCCCCGACC ATGAAATTGT GTTGGGAGCA TATACGAACG CCACTCATCG CCGCAACGGA
                                                                                   540
       GCTTTTGCCA CAGAGCCGGT ACCGGGGAGT GAGCTTATTA TGGATTATGA AGTGTCTCGC
                                                                                   600
60
       GGAGGGACTT TGCCTGACAT CAAGATCTCC GGTGCGGGTT ATATATTCGA CAAAGTCGGC
                                                                                   660
       GGACGCCCCG TAACGGATAA CCATTACGGG ATCGGTGAGG ACGATTCCGA TTCGGATTGC
                                                                                   720
       GAGATCAACA TCAATTGTCC TGAAGGTGCA GACTGGCAGG CAGAGAAGAA CGGTGTGGTG
                                                                                   780
       CAAATGATCA TGGTAAAAGG ACAGTATATC TCAATGTGCT CAGGCAACCT GCTCAATAAT ACGAAAGGAG ACTTTACTCC GCTGATCATT TCTGCCGGAC ACTGTGCTTC CATAACAACC
                                                                                   840
                                                                                   900
      AATTTCGGTG TAACGCAATC CGAGTTGGAT AAGTGGATCT TCACTTTCCA CTATGAAAAA AGAGGATGCA GCAATGGTAC ATTGGCCATC TTCCGTGGCA ACAGTATCAT CGGAGCTTCC
65
                                                                                   960
                                                                                  1020
      ATGAAGGCTT TCCTCCCGAT CAAAGGTAAA TCCGATGGTC TCTTGCTGCA ACTCAACGAT GAAGTCCCTC TGCGCTATCG TGTCTATTAC AATGGATGGG ACAGTACGCC CGATATTCCC
                                                                                  1080
                                                                                  1140
       TCGAGCGGTG CCGGTATTCA TCATCCGGCC GGAGATGCCA TGAAGATTTC CATCCTAAAG
                                                                                  1200
70
       AAGACTCCGG CTCTGAATAC ATGGATCTCC TCCAGTGGTT CCGGAGGGAC TGACGATCAC
                                                                                  1260
       TTCTATTTCA AATACGATCA AGGTGGTACG GAAGGAGGAT CGTCCGGTTC TTCTCTCTTC
                                                                                  1320
       AATCAGAATA AGCACGTGGT CGGCACACTG ACCGGAGGTG CCGGCAATTG TGGCGGGACG
                                                                                  1380
       GAGTTCTACG GCAGACTGAA CAGTCATTGG AACGAGTATG CATCCGATGG CAATACGAGC
                                                                                  1440
       CGCATGGACA TCTATCTGGA TCCCCAAAAC AATGGCCAGA CGACCATCCT CAACGGAACG
                                                                                  1500
75
      TATCGTGACG GTTATAAGCC TTTGCCCTCT GTGCCCCGGC TATTGTTGCA GTCTACAGGC
                                                                                  1560
```

```
GATCAGGTCG AATTGAATTG GACGGCTGTT CCTGCCGATC AATATCCATC ATCTTATCAG
                                                                                                          1620
          GTCGAATACC ACATATTCCG AAATGGAAAG GAAATAGCTA CGACAAAGGA GTTGTCCTAT
                                                                                                          1680
         TCGGATGCCA TCGACGAAAG TATTATCGGT AGCGGTATCA TTCGATACGA AGTAAGCGCA
CGCTTCATTT ATCCCTCGCC GTTGGATGGA GTGGAATCTT ATAAGGATAC GGACAAGACT
                                                                                                          1740
                                                                                                          1800
         TCTGCCGACC TTGCCATAGG AGACATTCAG ACCAAGCTGA AGCCGGACGT AACACCTCTC CCCGGAGGAG GAGTATCATT AAGCTGGAAA GTTCCTTTCT TAAGCCAGTT GGTTTCCCGA
   5
                                                                                                          1860
                                                                                                          1920
         TTCGGAGAAA GCCCCAATCC TGTGTTCAAA ACCTTTGAAG TGCCCTATGT TTCTGCCGCA
GCCGCACAAA CCCCCAATCC TCCCGTTGGC GTAGTCATTG CAGACAAGTT TATGGCCGGT
ACATATCCCG AAAAGGCTGC TATCGCTGCC GTTTATGTAA TGCCATCCGC TCCGGACTCT
                                                                                                          1980
                                                                                                          2040
                                                                                                          2100
 10
         ACTITICCACC TCTTCCTCAA GAGCAACACA AACAGAAGAT TGCAGAAGGT GACAACTCCC
                                                                                                          2160
         TCCGATTGGC AGGCCGGAAC ATGGTTGAGG ATCAATTTGG ATAAGCCGTT CCCGGTGAAT
                                                                                                          2220
         AATGACCATA TGCTTTTTGC CGGTATCAGA ATGCCTAATA AGTACAAGCT CAATCGTGCT
                                                                                                          2280
         ATCCGTTATG TAAGAAATCC GGATAACCTT TTCTCCATTA CCGGTAAGAA GATTTCATAT
                                                                                                          2340
         AACAACGGAG TCTCTTTCGA AGGCTACGGA ATACCCTCG TCTTGGGCTA TATGGCTATC
AAATATCTGG TGGTAAATAC CGATGCTCCG AAGATCGATA TGTCGCTGT ACAGGAGCCT
TATGCTAAGG GAACGATGT GGCTCCATTC CCCGAATTGG TCGCCTATA TGTCTATAAG
AACGGAACAT TTATCGGCAC ACAGGATCCA TCCGTCACAA CTTATTCGGT TTCAGACGGA
                                                                                                          2400
 15
                                                                                                          2460
                                                                                                          2520
                                                                                                          2580
         ACAGAGAGCG ATGAATACGA AATAAAACTG GTATATAAGG GATCGGGCAT TTCGAATGGC
                                                                                                          2640
         GTTGCTCAGA TTGAGAATAA CAATGCTGTC GTTGCATATC CGTCTGTTGT AACAGATCGT
TTCAGCATTA AGAACGCTCA TATGGTTCAC GCTGCCGCCC TCTACTCATT GGATGGCAAG
                                                                                                          2700
 20
                                                                                                          2760
         CAGGTTCGTT CTTGGAACAA CCTCCGCAAT GGCGTGACAT TCAGTGTTCA AGGACTTACG
                                                                                                          2820
         GCCGGTACTT ATATGCTCGT TATGCAGACG GCAAACGGCC CTGTGAGCCA AAAGATCGTG
                                                                                                          2880
         AAGCAG
                                                                                                          2886
 25
         (2) INFORMATION FOR SEQ ID NO:73
                (i) SEQUENCE CHARACTERISTICS:
                       (A) LENGTH: 2106 base pairs
 30
                       (B) TYPE: nucleic acid
                       (C) STRANDEDNESS: double
                       (D) TOPOLOGY: circular
               (ii) MOLECULE TYPE: DNA (genomic)
 35
              (iii) HYPOTHETICAL: NO
               (iv) AMTI-SEMSE: NO
 40
               (vi) ORIGINAL SOURCE:
                       (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
               (ix) FEATURE:
                       (A) NAME/KEY: misc_feature
45
                       (B) LOCATION 1...2\overline{1}06
               (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73
         TGGAAAAAA GTAATCCTCA ATACCGGGGC AGGACAGCAA ATTACCGAGA TCATTATAAT
                                                                                                            60
50
         ACAGAAAGCT ATTTAAGCCT AAGGAGGAAT CAAACAATGA AATATCTTAT CAGACTCTTC
                                                                                                           120
         TTATCATTGA TGTTACTCTC TCTCTGGACG GGCTGTACAC ACGAGGAGCT CTCTATTTGC
                                                                                                           180
         GATGGCGAGA ATACGCTTGT TTTACGCGTA GAGACCGGTA AAGCCCCAAA TGCTCGTGCC
                                                                                                           240
         ACAGAACCCG GTCAGGGCAT ATACAATGAG AATAAAGTAG GCTCCATTTC TGTGCTCTTC
                                                                                                           300
         TATTTAGAGG GACAACTTCG TTGGCAGGTG AAGTCTACAG ACTATCAAAT CCATGAAGGG
                                                                                                           360
        GCCTATATCA TTCCGGTCAA AGAGCAAATG CGACCACTAT TCAATGGCAA CAACAACTTC AGCATTTG TAGTGGCCAA TCTCGATTTC AATGCTCCGG CCACAGAAGC TGCGCTTTCT CAATTTGTGG TAGAGAAATC TATTGAAGTC TCTTCTACGA CAGCCCTGC CGATTTCGTA
55
                                                                                                           420
                                                                                                           480
                                                                                                          540
        ATGCTTGCTC ATGGCAATAA GCAGATCAAT ATGGCTACGA CAGAAGGGAA ACTGTTGGGG
                                                                                                          600
        GATTATAAAC TCAAACGAGT GGCAGCAAAG ATTCGCATGA TAAAACCCAC CATCAATGTG
                                                                                                          660
60
        CAAGGATATG AAGTGGTCGG AAATATACAG GCAAAGTTTC GCAATTCGGT AACGAAGGGG
                                                                                                          720
         TTCCTTACCA CAGAAGCTCA AGAGATCCCA GCTGCTGCAT CCTATAAGAC ATCGGAATAT
                                                                                                          780
        CTTGATATTG CAGAGTCGGC ACCTGCCAAT TCTATCCATT TCTATTCTTA CTATAACAAA
TGGACACTCT CCACACCGGA GAAGCGACCG GAATTCTTCA TCATGGTCAA ATTCAAAAAG
                                                                                                          840
                                                                                                          900
        ACAGGACAGE CGGACAACAC AGCCAAACCG TACTACTACA GAGTGCCCT CGAATCTCAG
GACAATCAG TCAAGAGCAA TGTCCTCTAT AATCTGAATG TGAAAATCGA AATCTTGGGT
TCTTTACAAG AGCCGAAGC TGTTTCTGTA AACGGCACAC TCGCAATAGA AGAATGGATT
CTCCATCAGG ATGCATTCAA TCTGCCTGCC ACCAATTACT TGATAGTGGA ACAGCACAAA
                                                                                                          960
65
                                                                                                         1020
                                                                                                         1080
                                                                                                         1140
        ATCTTCATGA ATAACGTGAA CACATACTCG GTGAAATATC AAACTTCGCA GAAACCAATC
AGCATTAGCA TACAGTCAGT TACCTTTAGC TACGTCTCTT CTGATGGCAC TCAGCACAAT
                                                                                                         1200
                                                                                                         1260
70
        GATCTTGTAG CAAGTAGTAG CGACCAGTAT CCTACGATTA CAAGCGATAA TACAAGCATC
ATAATCACTT CCAAGATACC GGTTAATAAC GTACCAAAGA AGATCGTTTT TGAGGTAACT
                                                                                                         1320
                                                                                                         1380
        AATGGGGTAG CCGGTTTGAA AGAGACTGTC ACAGTACTCC AATATCCTGC ACAATTTATT
                                                                                                         1440
        GTCAATACAC TTGGCACAGC ATCGGCATGG AGACCAGACG GATCTTTGGC TCCGGGGCTT
                                                                                                         1500
        AACAATAAAG CGATTTACCA TGTCGTAGTA CTGGTTCCAC CCGAGAATTT ATTTGAAGAT
                                                                                                         1560
75
        GGGACACAGA CAATCATCGG TTATCCCCCC ACTGAAACAA TTTCTTTTCA TAAGAAAGAG
```

```
AACAATACCT ATCCGATAGT ATGGTCTGAC ACAAATACGA CAAAACAGGA CCTTGAGACA
        TCAAGAATGA TTTCACCTTC CTTTGAGTTA GCCTCCCAAC TTGGGGCTAC TCTCCCGATG
                                                                                   1740
        CCCTATCTCG AGTATTGGCC AGGGACATCA TATCTCCTTG ACTATTCGGG AAACTATAAT
                                                                                   1800
        5
        GAAGAAATTA AATTCGATGA CTGGCGTTTG CCGACAGAAG CTGAGATCAA ATTGATAGAT
                                                                                   1920
        AAGCTGCAAC ATAATGAGCA GAGTGCTGTC CAAGCTATCA TGACAGGGAA TTATTATTGG
                                                                                   1980
        GATAGTTACT CTGCAAATGG GTCTTATAAA ATGCAAGGAG GAGGGGGCCA AGGAAATTCC
                                                                                   2040
        TCCAAAGCCT ATGTTCGTTG CGTGCGGGAT GTGAAAAAGC CGATTCGTGA CAAGAAGTCA
                                                                                   2100
        GGTAAG
                                                                                   2106
 10
        (2) INFORMATION FOR SEQ ID NO:74
             (i) SEQUENCE CHARACTERISTICS:
15
                   (A) LENGTH: 3936 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDHESS: double
                   (D) TOPOLOGY: circular
20
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
            (iv) ANTI-SEUSE: NO
25
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROHOMAS GINGIVALIS
            (ix) FEATURE:
30
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...3936
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:74
35
       AAGAGGGTAG ATAGCTATGA ATGCAGACAA AAAGCTTGCA AATGTGCAAT TTGTGTGATA
                                                                                     60
       CAAAAATTTA CTAATGTAAA ACTAAATGAT ATGCGAAAAA TTTTGAGCTT TTTGATGATG
                                                                                    120
       TGCTCTCTGC ATTTAGGTCT ACAATCTCAG ACTTGGCATG GAGATCCGGA CTCAGTGGCA
                                                                                    180
       GCCCTACCTT CTATCGGTAT TCAAGAGTCA AGTTGTACCC GAATCACGTT CGAGGTTGTT
                                                                                    240
       TTCCCCGGAT TTTATAGTGT GGAAAAACGA GAAGGCAACC AAGTCTTTCA GCGCATTTCC
                                                                                    300
40
       ATGCCGGGTT GTGGCTCGTT TGGGAATCTG GGCGAAGCTG AATTGCCTGT TTTGAAAAAG
                                                                                    360
       ATGATAGCCG TTCCGGAATT TTCAACAGCT AACGTTGCTG TAAAAATCAA AGAGACGGAG
ACATTCGACA ATTATAATAT CTATCCTAAT CCTACCTATG TCGTAGAGGA GTTGCCTGAG
                                                                                    420
                                                                                    480
       GGGGGGACTT ATCTGGTAGA GGCTTTCGCG ATAAACAATG ACTATTATAG CCAAAATGTA
                                                                                    540
       AGCCTCCCTT CTACTCACTA TGTCTATTCT CAAGACGGGT ATTTTCGCTC ACAAAGATTT
                                                                                    600
45
       ATCGAAGTTA CCCTGTATCC TTTTCGATAC AACCCTGTCC GACAAGAAAT TCTATTTGCA
AAAAAAATCG AGGTTACAAT AACTTTCGAT AATCCTCAGC CACCTTTACA AAAAAACACC
                                                                                    660
                                                                                    720
       GGCATATTTA ACAAAGTAGC CTCCTCTGCA TTTATTAATT ATGAAGCTGA TGGCAAATCG
       GCGATAGAAA ATGATATGGT GTTCAGTCGT GGTACAACAA CGTACATAAG CGGAAATGTT
                                                                                    840
       GCCAGCAACC TCCCTCAGAA CTGTGACTAC TTGGTTATTT ACGATGATAT GTTCAACGTA
                                                                                   900
50
       AATCAACAAC CACACGACGA AATCAAACGG CTGTGCGAAC ATAGAGCCTT CTACAACGGC
                                                                                   960
       TTTGATGTAG CTGCTGTAAG TATAAAGGAC GTATTGAATA GCTTCCCATC AAATGCCACC
                                                                                  1020
       TCATACATCA ACGAAACTAA ACTGAAAAAT TTCATTCGCT CAGTTTACAA CCAAAGCAAT
                                                                                  1080
       GCGAAGAGGA CTTTAGATGG CAAACTGGGA TACGTGCTAC TGATCGGAAA ACCATTGAGC
                                                                                  1140
       AAATATTTGG CTGACACTGA TAATACAAAA GTCCCAACCT CTTTTATTCA TAATGTCTCC TTAATTCCAA GTCATCCAAC TTTTGGTTCC ATATGCGCCT CCGACTATTT TTTTAGTTGT
                                                                                  1200
55
                                                                                  1260
       GTTTCGCCCC TTGATACTGT CGGCGATTTG TTTATCGGTC GATTTAGCGT CACCAATGCT
                                                                                  1320
       CATGAATTGC ACAATCTGAT TGAAAAGACT ATCAACAAG AAATCTCATA TAATCCTATT
                                                                                  1380
       GCACACAAAA ATATTCTTTA CGCAGAAGGG AAAGGCTGCG ATGCTCCAAT CTTACGTTTA
       TTCTTAAAAG AAATCGCCTC TGGTTACACA GTCAACTCTA TCTTAAAATC TAATCAGGTC
                                                                                  1500
60
      TCTGCAATAG ACTCGATATT TGACTGCTTG AATAATGGTT CCCATCATTT TTATTTTAAC ACTCATGGAA TGCCGACTGT TTGGGGGATA GGGCAGGGAC TCGACGTCAA TACTCTAACA
                                                                                  1560
                                                                                  1620
       GCCCGATTGA ACAATACATC TTCGCAGGGA TTATGTACGA GTCTATCATG TAGTTCGGCT
                                                                                  1680
       GTAGCAGATT CAACTATTAG ATCGCTTGGA GAAGTCCTGA CCACATACGC ACCTAACAAG
                                                                                  1740
      GGATTCTCGG CTTTCTTAGG AGGAAGCAGA GCCACCCAAT ATGCCGTTTA TTTAGAAGGC
CCCTGTCCTC CGTCAGAATT TTATGAATAT TTACCTTATT CTTTATATCA CAATCTCTCG
                                                                                  1800
65
                                                                                  1860
       ACTGTTGTTG GCGAAATGTT GCTATCATCC ATTATCAATA CTAATTCTGT TGATACGTAT
                                                                                  1920
       TCGAAATTCA ACTTCAATTT GCTTGGCGAC CCTGCACTAA ACATTATGGC TCATGGCATG
                                                                                  1980
      GAGGTTAGTA ATTGTATTAC ACTACCAAAC AACACCATTA TAAGCAGTCC GATAACAATA
                                                                                  2040
      AAAAATGGTG GCTGCCTAAA AATACCGGAA AAAGGAGTTT TGCATTTTAC TAATAATGGC
70
      TCCATACAAG TCATGTCCGG AGGAACTCTG GAAATAGGCA ATCAGGCTAA AATATCCGGA
                                                                                  2160
      GAGACCGGTG CTAACCCCAC CTTTATTACC GTTTACGGCG ATGGTCTTGC GATTAACAAG
                                                                                  2220
      CAGGTAGAGA TAGACAATAT AGACCGACTT AACTTGTTTT CTACGCATTC GGTCATGCCC
                                                                                  2280
      AAATTTCATT TTGACAGTGT GAAATTCAAC AGTGCCCCGC TGTATACAAC GAACTGTATT
                                                                                  2340
      GTGGAGATAA GCAATTGCGA ATTTACCAAT CGAAGTGACA TTATTTCAAA GAATTGTGAC
                                                                                  2400
75
      CTAAGCGTTG AAAACAGTAT GTTTAGCAGT TCGGGGATAA CGGTATTCAA GCCTATGGCT
                                                                                  2460
```

```
ACAAGCTCCA TCACCGGATT ATCTACAAAA GCAAAGATTA CCGACAATAC TTTTTTTGCG
ACAGGAAACT TCGCCTACCA TATCACAAAC ACGCCAGGCT TAACAGCAAC CTCCAATGCT
                                                                                                   2520
                                                                                                   2580
        GCCATCAAGT TAGACAATAT TCCTGAGTAT TACATTTCCG GTAATAAAAT AGTCAATTGC
GATGAGGCTC TTGTACTAAA TAATAGTGGC AACAGAACGA ACAGACTCCA CAATATCACA
                                                                                                   2640
                                                                                                   2700
  5
        CGGAATGTGA TAAAAAACTG TAGGATTGGG AGCACGCTTT ATAATTCCTA TGGTATTTAC
                                                                                                   2760
        AACCGAAATA AGATCAGTAA CAATCATATA GGAGTACGTC TCCTCAACAA CAGTTGTTTT
                                                                                                   2820
        TATTTCGATA ATGCTCCTGT AATCAATGAA GAAGATAAGC AGACGTTTAT TTCTAATAGG
                                                                                                   2880
        ACTTGGCAGC TCTATTCATC AAACGGTACA TTCCCTCTCA ACTTCCATTA CAACAGCTTG
CAGGGGGGAG ATACAGATAC ATGGATTTAC AACGACACGT ATACGAATCG CTATATTGAC
                                                                                                   2940
                                                                                                   3000
        GTTTCAAATA ATCACTGGGG CAACAATGAT TTGTTTGATC CGAATCAGGT TTTCAATACG
CCAGACTTGT TCATTTGGAT ACCTTTTTGG GATGGATTGC CAAATGGGA ATCGGGCAAT
10
                                                                                                   3060
                                                                                                  3120
        AGCTCTGCTG AAGCAGTAGA ATTCCAAACA GCATTGGACT GTATTGGCAA TAGCGATTAT
                                                                                                   3180
        CTTTCGGCAA AAGTGGCTCT CAAGATGATG GTTGAAACCT ACCCGGAATC CGACTTTGCA
                                                                                                  3240
        ATAGCTGCTT TGAAGGAATT GTTCAGGATA GAGAAAATGT CAGGCAACGA TTACGAAGGC
TTGAAAGATT ATTTCAGATC CAATCCAACC ATCATCTCTT CCCAGAACTT GTTCCCGACA
                                                                                                  3300
15
                                                                                                  3360
        GCTGATTTCC TGTCTGCGCG ATGCGATATT GTGTGTGAAA ACTATCAGTC TGCCATCGAT
TGGTACGAAA ATCGCTTGAA TAGTGAAATC TCCTATCAGG ACAGTGTTTT TGCAGTCATT
                                                                                                  3420
                                                                                                  3480
        GACCTTGGTG ACATTTATTG GAATATGCAG TTAGACTCAC TCAGAGGGAC TGGTATAGAT
TTGAACATAC TTTCCTGTGA ACAAAGGAAA TCGCTCGAAA GCCATCAAAA TGTAAAAAAT
                                                                                                  3540
                                                                                                  3600
20
        TATTTGTTGT CAACTCTTCC CGAATCAACA GGTACTCTCC TGCCTCCATT AGAATGCAAC
                                                                                                  3660
        AAATCAAGCC TTGATAAATC CAAGATAATC TCTATTTCGC CCAATCCGGC GAAAGCTGTT
GTAACAATAA TCTACTATAC CGATAACCCT TCCTGTTCTG TAATAAAAAT ATATGGAATA
                                                                                                  3720
                                                                                                  3780
        AATGGAGCCT CGGCTGATAT AACCGGGTTG CCCAAACATC TATCCGAAGG TTATTACAGC ATACAGTTCA ATACATCCAA CTTTGATCCC GGTTTCTACC TGGTAACGCT AAATGTTGAT
                                                                                                  3840
                                                                                                  3900
25
        CAGAAATTA TAGATACGGA AAAATTACGA ATCAAA
                                                                                                  3936
        (2) INFORMATION FOR SEQ ID NO:75
30
               (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 2814 base pairs
                     (B) TYPE: nucleic acid
                     (C) STRANDEDNESS: double
                     (D) TOPOLOGY: circular
35
             (ii) NOLECULE TYPE: DNA (genomic)
            (iii) HYPOTHETICAL: NO
40
             (iv) ANTI-SENSE: NO
             (vi) ORIGINAL SOURCE:
                     (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
             (ix) FEATURE:
                     (A) NAME/KEY: misc feature
                     (B) LOCATION 1...2814
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75
50
        TCAGAAAATT ATAGATACGG AAAAATTACG AATCAAATAA TGGCTATCAT GATGAAAAGT
        ATTGTTTTTA GAGCATTTCT AACGATTTTG CTCTCGTGGG CAGCGATCAC GAATCCGACT
                                                                                                   120
        GCTCAAGAGA TCTCAGGCAT GAATGCATCC TGTCTGGCTG CTCCGGCTCA ACCGGATACT
                                                                                                   180
        ATCTTATATG AAAGTTTTGA GAATGGACCT GTTCCCAATG GCTGGCTTGA GATAGATGCT
55
        GATGCTGATG GTGCCACTTG GGGAAGCCCA TCAGGCTCTT TCTCTGTACC TTACGGACAC AATGGCCTTT GCACCTACTC CCATATACGT TCCGGTATCT CAACAGCGGG CAACTATCTG
                                                                                                   300
                                                                                                   360
        ATTACACCCA ATATAGAAGG AGCCAAACGG GTCAAGTACT GGGTATGCAA TCAGTATAGT
                                                                                                   420
        ACCAATCCGG AACATTACGC AGTAATGGTA TCGACAACGG GGACTGCCAT TGAAGACTTT
                                                                                                   480
       GTTTTGTTGT TTGATGATTC CATAACAGGG AAACCGACTC CTCTTGTATG GCGTAGACGA
ATCGTGGACT TACCGGAAGG GACCAAATAT ATTGCATGGC GACATTACAA AGTCACCGAC
                                                                                                   540
60
                                                                                                   600
        TCACACAG AATTCTTGAA ATTGGATGAT GTCACTGTGT ATAGGTCGAT CGAAGGGCCC
                                                                                                   660
        GAACCTGCTA CCGACTTCAC AGTAATCAAT ATTGGTCAGA ATGTGGGACG ATTGACTTGG
                                                                                                   720
        AACTATCCGG AGGATTATCA ACCGGAAGGA AAGGGGAATG AAGAGTTGCA GCTTAGCGGC
                                                                                                   780
       TACAACATCT ATGCGAACGG TACACTACTG GCACAAATAA AAGATGTCTC CATACTGGAG
TATGTGGACA GCACTTACTC TTTGCGAGAC AATCCCTTGC AAGTGGAGTA CTGCGTTACA
                                                                                                   840
65
                                                                                                   900
       GCCGTTTACG ATGAAAGCAT AGAATCTTCG ACCGTATGTG GCACGCTGCA TTACGCCACG
GATGCCATCC TTTATGAAAA TTTTGAGAAT GGACCTGTTC CCAATGGTTG GCTTGTGATA
                                                                                                   960
                                                                                                 1026
       GACGCTGATG GAGATGGATT TAGCTGGGGA CACTATTTGA ATGCATACGA CGCTTTTCCC
                                                                                                 1080
       GGCCATAATG GAGGCCATTG CTCCTTGTCG GCTTCTTATG TTCCGGGTAT AGGCCCGGTG
                                                                                                 1140
70
       ACTCCCGACA ACTATCTGAT TACCCCCAAG GTTGAAGGAG CCAAACGTGT CAAGTACTGG
                                                                                                 1200
       GTAAGCACGC AGGATGCCAA TTGGGCAGCG GAACATTACG CGGTGATGGC TTCGACAACG
                                                                                                 1260
       GGGACTGCTG TCGGAGATTT CGTCATATTG TTCGAAGAAA CCATGACAGC GAAGCCGACC
                                                                                                 1320
       GGCGCATGGT ATGAAAGAAC CATCAACTTA CCTGAAGGGA CTAAATACAT CGCATGGCGG
                                                                                                 1380
       CATTACAACT GTACCGATAT ATATTTCTTG AAGTTGGACG ATATCACTGT ATTCGGGACT
                                                                                                 1440
75
       CCTGCATCAG AGCCCGAACC TGTTACCGAT TTCGTTGTCT CGCTTATTGA AAACAACAAG
                                                                                                 1500
```

```
GGACGATTAA AGTGGAATTA TCCTAACGGC TACGAACCCG ATAAGACTGA TGATAAAGAC
        CCATTGCAGC TTGCCGGCTA CAATATCTAT GCAAACGGCT CGCTCCTTGT TCACATACAA
GACCCGACTG TTTTGGAGTA TATCGATGAG ACTTATTCTT CACGAGACGA TCAGGTGGAA
                                                                                               1620
                                                                                               1680
         GTGGAATATT GTGTCACTGC CGTTTATAAC GACAATATCG AGTCCCAATC GGTTTGCGAT
                                                                                               1740
  5
         AAGCTGATTT ATGATTCTCA ATCGGACATT ATCTTATATG AAGGCTTTGA GGCCGGAAGT
                                                                                               1800
         ATTCCTGAAG GCTGGTTGTT GATTGATGCT GATGGCGACA ATGTTAATTG GGACTATTAT
                                                                                               1860
         CCTTGGACTA TGTATGGACA TGACAGTGAG AAGTGTATTG CATCCCCTTC GTACTTACCG
                                                                                               1920
         ATGATTGGCG TTTTAACTCC GGATAACTAT TTGGTTACAC CCAGACTCGA AGGAGCCAAG
                                                                                               1980
        CTTGTCAAGT ATTGGGTAAG TGCGCAAGAT GCTGTTTATT CGGCTGAGCA TTATGCTGTG
ATGGTTTCTA CTACGGGAAC TGCTGTTGAA GATTTTGTCC TCTTGTTCGA AGAGACAATG
                                                                                               2040
 10
                                                                                               2100
         ACCGCTAAGG CTAACGGTGC ATGGTATGAG CGAACTATTA CATTGCCTGC AGGAACAAAA
                                                                                               2160
         TATATTGCCT GGCGGCATTA TGATTGCACC GATATGTTTT TCTTGCTCTT GGATGACATT
                                                                                               2220
        ACGGTTTATC GTTCTACTGA GACTGTTCCC GAGCCTGTTA CTGATTTCGT TGTCTCGCTT
        ATTGAGAATA ACAAGGGTCG CCTGAAATGG AATTATCCTA ACGGCTACGA ACCCGATAAG
                                                                                               2340
        ACTGATGATA AAAAACCATT GCAGCTTACC GGCTACAACA TCTATGCAAA TGGCTCGCTC
CTTGTTCACA TACAAGACCC GACTGTTTTG GAGTATATCG ATGAGACTTA TTCTTCACGA
15
                                                                                               2400
                                                                                               2460
        GACGGTCAGG TGGAAATGGA ATATTGTGTC ACTGCCGTTT ATAACGACAA TATCGAGTCC
                                                                                               2520
        CAATCGGTTT GCGATAAGCT GAACTATACT ATCACATCCT TGGATAATAT TCAATCTGAT ACAAGCTTGA AAATATATCC TAATCCGGCA TCGTATGTGG TAAGGATAGA GGGATTGAGT
                                                                                               2580
                                                                                              2640
20
        CGGAGCAAGT CGACAATCGA GTTGTATAAT GCGCTGGGAA TTTGCATATT AAGGGAAGAG
                                                                                              2700
        ACTCATTCAG AGAAAACGGA AATCGATGTT TCACGTCTCA ATGACGGAGT CTACTTGATT
                                                                                              2760
        AAAGTAGTCG GTGGAAATAA AACAACAACC GAAAAGGTAG AGATAAAGAG GCCG
                                                                                              2814
25
        (2) INFORMATION FOR SEQ ID NO:76
               (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 1818 base pairs
                     (B) TYPE: nucleic acid
30
                     (C) STRANDEDNESS: double
                     (D) TOPOLOGY: circular
             (ii) MOLECULE TYPE: DNA (genomic)
35
            (iii) HYPOTHETICAL: NO
             (iv) AHTI-SEHSE: NO
             (vi) ORIGINAL SOURCE:
40
                     (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
             (ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...1818
45
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76
        ATAATCTTCT GTACGATTCA TCACTCTGAG TTGGAAATAA TGAACAGCAT CATGAAATAT
        CAATTATATA CGGCCGTCAT AATGGCTCTC TCTGTATCAT CCGTTTGCGG TCAAACCCA
                                                                                               120
50
        CGAAATACAG AAACCAAACG CCCCGACACG CTGCGCAGGG AGCTTACTAT CGTTAATGAC
                                                                                               180
        CAGACTGTGG AGATGGAGCA TGCGGATCCG CTTCCGGCTG CATACAAGGC CATCGAACCT
        CGATTAAAAC CTTTCCGTCC GGAATATAAC AAGCGTACAT TCGGATTTGT CCCTGAAGTT
        TCCTCTTCAG GCAGGAACAA TCTTCCGAAT ATCCTGCCGA CGGAAGGTCA TATGAAGCAC
                                                                                               360
       CGGGGGTACC TGAATATCGG TATCGGCCAT ACGCTAAACC AGCGAATGGA TGCCGGCTAT CGTCTGATAG ATGCAGAGCA GGAGAGACTG AATCTTTTCC TCTCCTATCG TGGGATGAAA
                                                                                               420
55
                                                                                               480
        TCGGCTTTCA ATACCGGTGA CTTCGACGGC GACAGAAAGG ATAGACGAAT GATGGCAGGA
                                                                                               540
        GTGGACTACG AGCAGCGCAG GCCTTCCTTT GTGCTTGCTA CCGGCTTGTA TTATTCGAAC
                                                                                               600
        CATTATTTCA ATAACTACGG ACGGGGAGCT ACCACCAATG TGGGCAGCAT CCCTCAGCTA
                                                                                               660
       TCGACACCTG TTACTCCTCA GATGGACAAC GGGACCCACA ACGTCCGTGT ATACTTGGGT GCAAAAAATG ATGTGATCGA TGCCAGGATC GACTATCGTT TCTTCCGTTC TATTCCCTAT
                                                                                               720
60
                                                                                               780
        840
        AGTAATGAGT TGTCCGATGA TATTAAGCTC GGTGTCGAAG TTCGTACGGG AGGATTGTTT
                                                                                               900
        TTTGCCAAAA ACAGCGAAAT GATTCAAACG GGCGTTCTGT CCGAAACCGA CCGCAACCTG
                                                                                               960
       TATTATGTGG AGGGCGCCC CACAATCGGA TTTGTCGGAG ACTCGGACAA TATGCAATGG
                                                                                              1020
       AACATACAGG CCGGAGTAGG GATTTCTTCC CATTTCGGAG CCAAAGGGAG GTTGTTTTC
TGGCCTAAAC TGGATGCTTC GCTTAGTATC TTCCCTTCAT GGCGTGTGTA TGCGAAAGCC
65
                                                                                              1140
       TTCGGCGGTG TGATTCGAAA TGGTCTCGCC GATGTTATGC AAGAGGAGAT GCCCTACCTG
ATGCCCAATA CGATTGTACT CCCTTCGCGC AATGCTTTGA CCGCCCAATT AGGGGTGAAG
                                                                                              1200
                                                                                              1260
       GGGAATATAG CCGATGTGGT ACGTATGGAG GTTTATGGCG ACTTCTCCAA GCTGACAGGT
GTGCCTTTCT ATACTCCGAC TCTACCCTTA TATAATCCAT CCGACTTGTA TCAGTATAAT
                                                                                              1320
70
                                                                                              1380
       GTGAGTTTCT TGCCGATATA TGCCGACGGC AGCCGCTGGC GCGCAGGTGG TAAGCTGGAA TACTCTTATC GCGATATGCT CCGCTTTCTG GTAGACGCAT CCTATGGCAA GTGGAATTTG GATGGAGGC TTGTCGCCTC CATGCAGCCC GATCTTATAT TGAAGGCAGA AGTAGGTGTT
                                                                                              1440
                                                                                             1500
                                                                                              1560
       CATCCCATTG CCCCATTGGA TGTCAGACTC CGGTATACAC AGCTGAACGG ACGGTATCGG
TATTCTTTCG GCTCGGCTGG CTCGGAAGCC TTGGGTATCG GTAATGTACA TCTTCTTAGT
                                                                                             1620
75
                                                                                              1680
```

| | GCGGATGTTT CATACAAGCT GAAAAAGAAC TTGAGCCTTT ATCTCAAAAT CGATAATATG CTGGCGGAAA CGACAGAGCT TATCGGTTAT TATCCTATGC AGCCGTTCCA TTGTTTCGCC GGTTTTAGCT GGACTTTC | 1740 1800 1818 |
|------------|---|-----------------------------------|
| 5 | (2) IMFORMATION FOR SEQ ID NO:77 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1071 base pairs (B) TYPE: nucleic acid (C) STRAIDEDNESS: double (D) TOPOLOGY: circular | |
| 15 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 20 | (∀i) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROHONAS GINGIVALIS | |
| 25 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11071 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77 | |
| 30 | AACTCCCATG CAACAGTCTC CGTTATATGT TCAATGATGG AAAAATGTAT CTTTGCTCAC TATCCACATA ACCTTGTGTT CATGATTCGC AAGCATTTCG GTATCATTTT GGGATTTCTT TCTCTTGTGT TTTCGGCAGG TGCTCAACAA GAGAAGCAGG TGTTTCATTT TCTGAACCTT CCGGCTACTG CACAGGCTTT GGCTGCCGGA GGCAAAGCTA TCACCATCGT AGACGACAAT | 60 120 180 240 |
| 35 | CCCGGACTGG CTTTTGAGAA TCCGGCTCTG CTCGGATATG AATCCGGTGG CCGCGCCTTT CTTTCCTATT TATATTATAT GAGTGGTTCG CATATGGGCA ATGCCTGTTA TGCCTCGTCC GTCGGAGAGC GTGGCATGTG GGGTGTTGGC ATGCGTTTC TGAACTACGG GTCTATGCAA GGATACGATC AGAATGCGAT TGCCACCGGC TCTTTTAGTG CTTCGGATAT AGCTGTACAA GGATTTTACA GCCATGAACT GAGCAACCAC TTCCGCGGTG GAGTCAGCCT AAAAGCATTG | 300 360 420 480 540 |
| 40 | TATTCTTCTA TCGAGACGTA TAGTTCCTTT GGCCTTGGTG TGGATGTCGG TATCAGTTAT TACGACGATG ACAAAGGATA TTCCGCTTCC GCTCTGTTCA AGAACGTAGG GGCGCAACTG AAAGGCTATA ATGAAGAACG GGAACCGCTC GATTGGGATT TCCAGCTCGG CTTTTCCCGC AGTTTTATCA ATGCTCCGTT TCGCTTGCAC ATCACGTTGT TCAATCTGAA TCCGCACTAT TCCAAGCGTC TTGTACCACG CGATCTGTCC AAGATGCAAA AGTTCCTCCG ACACTTCTCG | 600 660 720 780 |
| 45 | ATAGGAGCAG AATTTACTCC TTCCGAGAGG TTTTGGGTCG GGCTGGGATA TACGCCACAG ATTGCACAGG ATTTCGAGGT GGAAGGCGGC AACAAATGGG GAGGTCTTCC GGCCGGCGTC GGTTTCACTT CAGGTGTAGT ACGTGTAGGC GTATCTGCTG CCACCTATCA TCCTGCAGCT CTTTCGTTCA TGTGTTCGGT AGGTATCCGT TTGGACGATA AGAGCATCTT C | 840 900 960 1020 1071 |
| 5 0 | (2) INFORMATION FOR SEQ ID NO:78 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1011 base pairs | |
| 5 5 | (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 60 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 65 | <pre>(vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROHONAS GINGIVALIS</pre> | |
| | (ix) FEATURE: (A) MAME/KEY: misc_feature (B) LOCATION 11011 | |
| 70 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78 | |
| 75 | CCTCAGCCCG TCGGCCTTAA AGAAATAACC ATTAAACCCA TGTGCCTCGA ACCCATAATT GCTCCGATTT CATCCGAGTT GCTCGACAG GAGCTGACTG CCGATCGTTT TCTGCGATG ACAAACAAAG CCGGCAATGA GATCTATGTT TTTACGGCCG AAGAAGCTCC GCATTGCATG AAAGAAGTAG GCCGACTGCG AGAAGAAGCC TTTCGGCATT ATGGCGGAGG TACTGGCAAG | 60 120 180 240 |

```
GCGATCGATA TAGACGAGTT CGACACCATG CCCGGGAGCT ACAAACAGCT GATCGTATGG
                                                                              300
      GATCCGCAAA ACAAGGCTAT ACTCGGAGGC TACCGCTTTA TCTATGGGCG GGACGTTGCT
                                                                              360
      TTCGATACCG ATGGCAAGCC TTTGCTGGCA ACGGCAGAGA TGTTTCGCTT CAGTGATGCT
                                                                              420
      TTTTTGCACG ATTATCTCCC CTACACAGTC GAATTGGGAC GTTCGTTCGT GTCGCTCCAG
                                                                              480
 5
      TACCAATCGA CACGGATGGG CACAAAGGCC ATTTTTGTGC TGGACAATCT TTGGGACGGT
                                                                              540
      ATCGGAGCAC TCACTGTAGT CAATCCAGAG GCACTCTATT TCTATGGCAA GGTGACCATG
                                                                              600
      TACAAAGACT ATGATCGGCG AGCTCGCAAT CTGATCCTGT ATTTTCTTCG CAAGCACTTC
                                                                              660
      TCCGATCCGG AAGGCTTGGT CAAGCCTATT CATCCCCTAC CGATAGAGAT CAGTGCGGAG
                                                                              720
      GACGAAGCCT TGTTCTCCTC ATCCGACTTT GACACCAATT ACAAGACTCT CAATATAGAA
                                                                              780
10
      GTGCGCAAGC TGGGTATCAA TATCCCTCCT CTCGTGAGTG CATATATAGC TTTGTCTCCG
                                                                              840
      GAGATGCGTG TTTTCGGCAC TGCAGTGAAT GAGTCTTTCG GAGAGGTGGA GGAAACCGGC
                                                                              900
      ATATTCATTG CTGTGGGTAA GATCCTGGAA GAGAAAAAAC AACGGCACAT AGAGAGCTTC
                                                                              960
      ATCCTCAGCC GGAACGAAAA AAAAGGTCTC GACAGTAGCA ATGGCCGATC A
                                                                             1011
15
      (2) INFORMATION FOR SEQ ID NO:79
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1698 base pairs
20
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
25
         (iii) HYPOTHETICAL: NO
          (iv) AUTI-SENSE: NO
30
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
35
                 (B) LOCATION 1...1698
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79
      TACGACGGAG CGAGACTGGT CTATACCCTA TTTCGGAATA GAAACGACAT TCACCCTATG
                                                                               60
40
      AAAACCATTG TAAGATACAG CCGCCTTCCG GTCGCTCTCT TCTTTTGCCT TTTGGGAGCT
                                                                             120
      GTGCACTTGT CTGTCGAGGC GCAGATGCTC AATACTCCTT TCGAGCTGTC GGATCAGATC
                                                                             180
      GTCCTGTCTC CCACCGAAAG GCAATACAGG GAGATTTGTG TGCAAACGAA AGAAAAAAGG
                                                                              240
      GGGGCCGATC TTTTCCCGTT GAGCGATAAG CTGCGCGATT CGGCCTATGT TCGTTTCGGC
                                                                              300
      TEGGECTATG GEGATATTGE GGGCGACTAT CTTCCGTACA ACGGCAATAA CTACTCCTCG
                                                                              360
45
      CTCTCGCTCG AATCGGGTGG TCGCATCAGT GTCCGTAACT ATGGCACATT GCAGGGCAGT
                                                                              420
      GCTTCCTACT CACGTGGCAT GCACAAACGC ATCGGCTGGA ATGCTCTGCG CAACGCCGAA
                                                                              480
      GCCTACTATC CCTATTTGGT GTCCGATTCG ACCGGCGGAG ACTATCATTT CGAAGACTAT
                                                                              540
      CGGCTTGCCG GCTACTATTC TTTTCGCGCC GGCCGCTTGC CCCTCGGTAT AGGCTTCTCA
                                                                              600
      TACAGGGGCG AAGTTGCTTA TCGGCTGACC GATCCGCGTA CGACCAATAC GACCGGTGCA
50
      TTGGAGCTTT CTTGTGCTAC CTCTTTGACG CTGCCTCGAG AGAACAGGCT ATCGCTTTCG
                                                                              720
      GCTGCGTATC TCTATCATAG ACAACACCTC ACACAGTACA ACTGGCGTCC CGGGCAGCAG
      GACAAATTCT TCGTCAGCTA CGGTTTCGGT CAGGTGGATG TCAGCAACAG CCCTATCTGG
      TTCGGTATCT CCAGAATGAA CTACGTCAAC GGATGGAAGC TTAGCTCCCG TCTGGATACC
      CGTAGGGGCG ATGCCATCGG TCTCGACTAC AGCGGCTACT TCCTCGATAC CGAaGAGAGG
                                                                              960
55
      TCGTCCATCA ATCTCTTTGC TTTGCTTTAC AATCGCCTGC GACTCTATGG TAGCTGGCAT
                                                                             1020
      CTGTCGGACT TCGATTTTC ATTTCAGCC GACTATGCTC TGCGCCAAGG GATAGAGCGG
                                                                             1080
      ATATACGAAG ACTACAAGCC GGATGATAAT TATCATATCT ACGACCTCCG TATCTTGGCC
                                                                             1140
      ATTCGCCGCT GGTATATGCT CAATGAGTTT TCTGCCCAAG CCCAAGCCTC CTACCGTATT
                                                                             1200
      CGCACGGATA GAGGTTGTGC CCTGAGAGTG AGTGCCGGTA GTGATTTCTA CGGCTATGAT
                                                                             1260
60
      GAGACGTATC GCAAGCATGG ACATCATACC ATGAGCGGAA TGCTACGTCC TTTTGCCGGT
                                                                             1320
      ATAGCCTATG ACCATGCCGG ATCCAAATTG GATTTTGGAC TTTCGCTTTC GGCTGCTTAT
                                                                             1380
      CGAATGGTGC TGACGCATTC GTATAAGATT CGTACCATCC AGAAAGAGCA GCTCGACTAT
                                                                             1440
      CAGCTGGCCT ATTTGCCCTA TGCCTATCGT AATAGAGAAG GCGTGGAGGT GCGTTCCTCT
                                                                             1500
      CTGTACGTCT CGATTCCGAT GCAGAATACC CACCGCCTGA TGACAGAGCT GCGGTTGTAT
                                                                             1560
65
      GGCGACCTGA TGAAAAGAAA GGACGGTATA GCCTATGGCA AAACGCCCGG TGTCATCTCA
                                                                             1620
      CATATCCTGT CCGATCCGCA AGCCGAACGA ACGTCCGGCC ATACCATCGG GGCTATCTGC
                                                                             1680
      AATATCTCCT ACCTCTTC
                                                                             1698
70
      (2) INFORMATION FOR SEQ ID NO:80
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2457 base pairs
                 (B) TYPE: nucleic acid
75
                 (C) STRANDEDNESS: double
```

```
(D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
  5
          (iii) HYPOTHETICAL: NO
           (iv) AHTI-SEHSE: NO
           (vi) ORIGINAL SOURCE:
10
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...2457
15
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80
       AGGACGAACG TTTTCTTATC TTTGTCCCAT AAAATTGGGA GAAGGGGTGC TTCCTGTAGC
                                                                               60
      AATCGGAATG CGTGGATGGC TGAGAACAAA CCCTCATCAC CTGAACCGGA TAATACCGGC
                                                                              120
20
      GTAGGAAACT CTCCGTCTGA CTATCTTCAC GGCGAAGCAA TCATTCCCCC TCTCTTCT
                                                                              180
       TTGTCCAACT TCAATGATAA GAGATTTATG AAAAAACTTC ACATGATTGC CGCCTTAGCC
                                                                              240
      GTCCTGCCTT TCTGCCTGAC GGCACAAGCA CCCGTCTCCA ACAGCGAGAT AGATAGTCTT
                                                                              300
      AGCAATGTGC AGCTCCAGAC CGTACAGGTC GTAGCTACTC GCGCCACGGC GAAAACCCCT
                                                                              360
      GTCGCTTACA CCAACGTTCG CAAGGCCGAA CTTTCCAAGT CCAATTATGG TCGTGACATC
                                                                              420
25
      CCCTATCTGC TGATGCTGAC TCCCTCCGTG GTAGCCACCA GCGATGCCGG TACGGGTATC
                                                                              480
      GGATATTCCG GCTTTCGCGT GCGTGGCACC GATGCCAATC GCATCAACAT AACTACCAAT
                                                                              540
       GGAGTACCCC TCAACGACTC CGAATCTCAG TCCGTCTTTT GGGTGAATAT GCCCGACTTC
                                                                              600
      GCCTCTTCCA TCGAAGACCT TCAGGTGCAG CGAGGTGTGG GTACTTCCAC CAATGGTGCC
                                                                              660
      GGAGCTTTTG GGGCAAGTGT CAATATGCGT ACGGATAATT TGGGACTGGC TCCTTATGGC
                                                                              720
30
      CGTGTCGATT TGAGCGGAGG TTCGTTCGGC ACATTCCGCC GATCGGTCAA ACTCGGTAGC
                                                                              780
      GGACGCATCG GTCGCCATTG GGCAGTGGAT GCCCGCCTGT CCAAAATCGG TTCGGACGGC
                                                                              840
      TACGTGGATA GAGGAAGCGT GGATCTGAAA TCCTATTTCG CACAGGTGGG CTATTTCGGT
                                                                              900
      AGCAACACGG CTCTCAGGTT CATCACTTTC GGAGGAAAAG AAGTTACGGG TATCGCATGG
                                                                              960
      AACGGTCTTT CCAAGGAGGA TGAAGCCAAA TATGGCCGCC GATACAACAG TGCCGGTCTT
                                                                             1020
35
      ATGTACGTGG ACGCGCAAGG AGTACCGCAC TACTACCACA ATACCGACAA TTACGAGCAG
                                                                             1080
      CGTCACTACC ATGCCATCAT GACGCACAGC TTCTCTCCTT CCGTTATCCT CAACCTCACG
                                                                             1140
      GCACACTACA CGGCCGGATA TGGCTATACG GACGAATATC GTACCGGACG TAAACTAAAG
                                                                             1200
      GAATATGCAC TGCAGCCCTA TGTGGAAAAC AGTGTGACCG TGAAGAAAAC GGATCTCATC
      CGTCAGAAGT ATCTGGACAA TGACTTCGGA GGACTCATCG GTTCGCTTAA CTGGCACACC
                                                                             1320
40
      GGTGCATGGG ATTTGCAGTT CGGGGCCTCG GGCAATATCT ATAAAGGAGA CCACTTCGGC
                                                                             1380
      CGTATCACTT ACATCAAAAA GTACAATCAG CCCTTAGCTC CCGACTTCGA ATATTATCGG
                                                                             1440
      AACAGGGCAG ACAAAAGAGA AGGTGCAGCC TTTGCCAAAG CCAACTGGCA GATCACTCCG
                                                                             1500
      GAACTGAACA TGTATGCCGA CCTCCAGTAT CGTACCATCG GCTACACGAT AAACGGCATC
                                                                             1560
      ACGGACGAAT ATGATGAGGT ACAGGGAAGT ATGCAGCACA TCGATTTGGA CAAGACCTTC
                                                                             1620
45
      CGCTTCCTCA ATCCGAAGGC CGGTCTTACC TATAGTTTCG ACGATGCTCA TACTGCCTAT
                                                                             1680
      GCTTCTGTTG CGGTAGCACA CCGCGAGCCT AACAGAACCA ATTACACCGA AGCCGGAATA
                                                                             1740
      GGACAGTATC CTACGCCTGA GCGACTGATC GACTATGAGC TGGGCTACCG CTATGCTTCG
                                                                             1800
      CCCCTCTTGT CGGCCGGAGT AGGTCTCTAT TATATGCAAT ACAAGGACCA ACTCGTGCTG
                                                                             1860
      GATGGCCGTT TGAGCGATGT GGGACAGATG CTCACAAGCA ACGTCCCCGA CAGCTACCGT
50
      ATGGGACTGG AGCTGACTCT CGGTTGGCAG ATCCTTCCTC GTTTGCTGCG TTGGGATGCT
                                                                             1980
      TCTTTCACTA TGAGTCGCAA CAAAATCGAC CGCTACGTAC AATATACATC CGTATATGAT
                                                                             2040
      GCGGACTACA ACTGGCTCGA ACTCAAGGAG GAGACCCTCG AAAGCACGGA TATAGCCTAC
                                                                             2100
      TEGECEAATG TEATTGEEGG CAGCATGETT ACCETETEE ATGEEGGTTT CGAAATGGET
                                                                             2160
      TGGACGAGCC GCTTCGTCAG CAAGCAATAT CTGGACAATA CACAGCGCAG CGATCGCATG
                                                                             2220
55
      CTTTCCTCCT ATTGGGTGAA CGACCTCCGC CTCGGCTATG TGCTGCCGGT TCACTTCGTT
                                                                             2280
      AAGAGAGTGG CACTGGGCGT ACAGCTCAAT AATCTCTTCA ACCTCATGTA TGCGTCCAAT
                                                                             2340
      GCCTACATCT ACGATGCCGG TTACGTACAG GCATCCGGAG AACTAAGTGC ATATGCCGAT
                                                                             2400
      CTGCGTTATT ATCCTCAGGC CGGATTTAAT GCACTGGGTA GTCTGACAAT CGATTTC
                                                                             2457
60
      (2) INFORMATION FOR SEQ ID NO:81
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1596 base pairs
65
                 (B)
                    TYPE: nucleic acid
                    STRANDEDNESS: double
                 (C)
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
70
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
75
          (vi) ORIGINAL SOURCE:
```

PCT/AU98/01023

```
(A) ORGANISM: PORYPHYROMONAS GINGIVALIS
             (ix) FEATURE:
                   (A) NAME/KEY: misc feature
  5
                   (B) LOCATION 1...1596
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81
        TACAGGCGAT CGGGAAGGGT TTGTCCTCGT ATCCTCCGGA ACAAACGTTC CTACGCAGAT
                                                                                      60
 10
        ACTCGGTTAC AGCCGAGAAG AGCGGTTCGA CTACGAGCCG GCACCAGAAC AAAGATGAAA
                                                                                     120
        AGAAGGTTTC TATCGCTGTT ACTGCTGTAC ATACTCTCTT CCATCAGCCT TTCTGCTCAG
                                                                                     180
        CGGTTTCCGA TGGTGCAGGG AATCGAGTTG GATACCGATT CGCTTTTCTC TCTGCCCAAG
                                                                                     240
        CGTCCTTGGC GCGCCATCGG TAAAACGATA GGCGTCAATC TGGCCGTATG GGGCTTCGAT
                                                                                     300
        CATTTCATCA TGAACGAGGA CTTTGCAGAC ATCAGTTGGC AGACTATCAA GAGCAATTTC
                                                                                     360
 15
        CAAACAGGCT TTGGCTGGGA CAATGACAAG TTTGTCACCA ACCTCTTCGC ACATCCTTAT
                                                                                     420
        CACGGATCGC TCTATTTCAA TGCAGCGAGG TCGAACGGTT TGAGCTTCAG GCACTCTGCT
                                                                                     480
        CCGTTTGCCT TCTTTGGCAG TCTCATGTGG GAGCTGCTTA TGGAAAACGA GCCACCGAGT
                                                                                     540
        ATCAACGACC TCTGTGCCAC CACCATAGGC GGTATAGCTT TGGGGGAGAT GGGGCACAGG
                                                                                     600
        CTGTCGGACC TGCTCATCGA CAATCGTACC ACAGGGTGGG AACGTATGGG GCGCGAGGTG
GCTATCGCTC TGATCAATCC GATGCGCTTT CTCAACCGTC TGACAGCAGG AGAGGTGACT
                                                                                     660
 20
                                                                                     720
        TCTGTCGGGA GTCGCAGCGG ACAGATATTT CAGTCTGTCC CCATAAACAT AGTCGTCGAT
GCCGGCTTTC GCTTTTTGGC AGACAAGCGG CATGCCCGAA CCGGTGCCAC GGCTCTGACC
                                                                                     780
                                                                                     840
        CTGAATCTGA GATTCGACTA CGGCGATCCA TTCCGAAGCG AGACTTTCTC TCCATACGAT
                                                                                     900
        TTCTTCCAAT TCAAAGCCGG ATTGAGTTTC TCCGAATCGC AACCTCTGCT GAGCCAGATC
                                                                                     960
 25
        AATCTGATCG GAATCCTAAG CGGATGCCAA CTGCTCGCAC ACGAACGAAC GGTTTTGGTG
                                                                                    1020
        GGAGGTCTCT TTCAGCACTT CGACTACTAC AATTCGGAAA AACGAATAAG CAAAAATTCG
GAGGAGGTAC TCGTCACCCC ATACCGTATC TCGCAAGTGG CAGCTCTGGG AGGCGGTCTT
                                                                                    1080
                                                                                    1140
        ATCTTCCAGC ACCACGGAAA ATTTCGACGA CGTCCTCTGG AGCTATATGC CGAGACCTAC
                                                                                    1200
        CTGAATGTCG TCCCGATGGG AGCCAGTCTG TCGGATCACT ACAACGTGGA CAATCGGGAC
                                                                                    1260
        TATAACCTCG GCAGCGGATT GAGCGGCAAG CTATACCTTG GTGCTACGTA CAATGATCTG
TGGAGCTGGC TCTTGGGAGT CGAAAGCTAT CGGCTCTACA CATGATCGG GTATGAAGAG
 30
                                                                                    1320
                                                                                    1380
        CCGCACCAGA AAAATACCGA TGTCAGCTCT TTTATGGTGC AGGGGGACGA AAGCAAGGCG
                                                                                    1440
        CGCCTACTGG TGACGAGTTC CGAGTTCGCA TTTCATCCTG GCCCCTGGCA TGTAGCCATC
                                                                                    1500
       GTCGCTCGCC GTTTCATCCG CAAAACAGCC TATCAATTCT ACCCTAACGT ATCATTCGAT
                                                                                    1560
35
       ACCGGCGACA TACAGCTGCG TGTCGGATTT CACTTC
                                                                                    1596
        (2) INFORMATION FOR SEQ ID NO:82
40
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 900 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
45
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
50
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
55
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...900
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82
60
       AAAAGAAGAA AAAAACAGAT GAAACGACTG ATTGTTTTTC TGGCAATGGG TGGCTTGCTG
                                                                                     60
       TTCACCCTTG CGAACGCACA AGAAGCAAAC ACTGCATCTG ACACTCCCAA AAAGGACTGG
                                                                                    120
       ACTATAAAAG GTGTGACCGG ACTAAATGCC TCTCAGACTT CTCTGACCAA CTGGGCTGCC
                                                                                    180
       GGTGGAGAAA ACACGGTGGC AGGTAACCTC TATTTGAACA TAGATGCCAA CTACCTGAAA
                                                                                    240
65
       GATAAATGGA GTTGGGACAA CGGTTTGCGT ACAGACTTCG GTCTGACCTA CACAACAGCC
                                                                                    300
       AACAAGTGGA ACAAAAGTGT AGACAAGATC GAACTCTTCA CGAAGGCCGG CTATGAGATC
                                                                                    360
       GGCAAACATT GGTACGGAAG TGCGCTTTTC ACTTTCCTCT CACAGTATGC CAAAGGATAT
                                                                                    420
       GAGAAGCCCT ÇGGATCACTT GACAGGAGTC AAGCATATCT CTAATTTCTT CGCTCCTGCA
                                                                                    480
       TATCTCACTC TCGGTATTGG TGCGGACTAT AAGCCCAATG AGAAGTTCTC TCTCTACCTC
                                                                                    540
70
       TCTCCTACAA CGGGCAAGCT GACTGTAGTA GCAGACGACT ACCTCTCAAG TTTGGGAGCC
                                                                                    600
       TTCGGGGTGA AAGTTGGTGA AAAGACAATG TTCGAACTTG GTGCTTTGGT AGTGGGTTCG
                                                                                    660
       GCCAATATAA ATCTGATGGA GAATGTCAAT TTGATAACCA AGGCTTCATT CTTCTCGGCT
                                                                                    720
       TATACGCACG ACTITGGCAA CATTGACATC AATTGGGAGG CTATGCTGGC CATGAAGATC
                                                                                    780
       AACAAGTTCC TCACGGCTAC GATAGCCACC AATCTTATCT ACGACGATGA TGTGAAGATC
                                                                                    840
75
       AACGATGGCC CGAAAATCCA GTTCAAAGAA GTTGTGGGCG TGGGTGTTGC GTACACTTTC
                                                                                    900
```

```
(2) INFORMATION FOR SEQ ID NO:83
 5
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 663 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
10
           (ii) HOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
15
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
20
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...663
           (xi) SEQUENCE DESCRIPTION: SEO ID NO:83
25
      ACGAGAGAGA GTGTGTTACA TTGTAGAACA AAACTCAAAA AAGAACGAAA AATGAAGAAA
                                                                                      60
      ATGATTTTGG CAGCTACTAT GCTGCTCGCA ACAATCGGTT TTGCAAATGC TCAGAGTCGT
                                                                                     120
       CCTGCTCTTA GACTGGATGC TAACTTTGTC GGTAGTAACT TAATGCAAAA AGTCGCAAAC
                                                                                     180
      ACGAGCGTGA ACAATAAGAT GATCGTAGGC TTACGTGTTG GTGCTGCTGC TGAGTTCGCT
                                                                                     240
30
       CTTAGCAATG ATGGATTCTA TCTCGCCCCC GGATTGGCCT ATACGATGAG AGGTGCTAAG
                                                                                     300
      ATGGAATCAC TAAGTGAAAC GACAACTCGC TTGCATTATC TGCAAATACC GGTGAATGCC
                                                                                     360
      GGTATGAGAT TTAGCTTTGC TGACAACATG GCTATTTCAT TGGAAGCAGG TCCCTATTTC
GCATATGGTG TCGCCGGAAC GATTAAGACT AAAGTTGCAG GCGTTACGGC TTCTGTAGAT
                                                                                     420
                                                                                     480
      GCCTTTGGTG ATAACGGATA TAACCGTTTC GACTTGGGCT TGGGCTTGTC TGCTGCCTTG
AGCTACGACC GTTATTACGT ACAAATTGGA TATGAGCATG GATTGCTTAA TATGTTGAAG
                                                                                     540
35
                                                                                     600
      GATGCTCCGG ATAAGACTTC TTTGCGTAAT CATGACTTCT TTGTGGGTCT CGGTGTTCGC
                                                                                     660
                                                                                     663
40
      (2) INFORMATION FOR SEQ ID NO:84
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 744 base pairs
                  (B) TYPE: nucleic acid
45
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) NOLECULE TYPE: DNA (genomic)
50
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
55
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...744
60
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84
      ATCAAACGAA TAGAAATGAA AAGGATTTTT ACTGTAGCCC TTGTGCTACT TGCTTCGGTC
      ACTATGGCCA TCGGACAAAG CCGCCCGGCA CTTCGCGTAG ATGCCAACTT CGTAGGCAGC
65
      AATCAGAGCA TGAAAAGAGA CGGATATGTG TGGGACACCA AAATGAATGT CGGCCTGCGG
                                                                                    180
      GTCGGTGCCG CTGCCGAATT CATGATCGGA TCAAGAGGAT TCTACTTGGC TCCGGGTCTG
                                                                                    240
      AACTATACGA TGAAGGGCTC CAAAACCGAA TGGGATATAC CCGAAATGGT TCCTGGTACC
                                                                                    300
      TATATTACGA TGGTTTCCAC TCGCTTGCAC TATCTGCAAC TGCCGATCAA TGCCGGCATG
                                                                                    360
      CGGTTCGACC TGATGAATGA CATGGCGGTT TCGATCGAAG CGGGTCCTTT CCTTGCATAC
                                                                                    420
70
      GGTATATATG GTACATATCG GCAGAAGTTG GAAGGATGGA AGCCGAACAA CTACAGCACA
                                                                                    480
      GAGTTTTTTG GCCCAACGCT TGGTGGCCCA ACAAATATCC GCTGGGACAT CGGGGCAAAC
ATAATAGCCG CATTCCACTA TAAGCGTTAT TATATACAGA TAGGCTATGA ACATGGATTT
                                                                                    540
                                                                                    600
      GTGGATATTG TGTCAGGTGG AGGTTCTGAT ATTCCCCGAC TGAACGACAA TAGGCAATCC
                                                                                    660
      TCTTCGACGA CCGCTCTAAG AGAAAAGGGA AATAACGAAT ACGCTTATAA TCGTGACTTC
                                                                                    720
75
      TTCGTGGGCA TAGGTTACCG CTTT
                                                                                    744
```

```
(2) INFORMATION FOR SEQ ID NO:85
  5
               (i) SEQUENCE CHARACTERISTICS:
                      (A) LENGTH: 633 base pairs
                      (B) TYPE: nucleic acid
                      (C) STRANDEDNESS: double
                     (D) TOPOLOGY: circular
 10
              (ii) MOLECULE TYPE: DNA (genomic)
            (iii) HYPOTHETICAL: NO
 15
             (iv) ANTI-SENSE: NO
             (vi) ORIGINAL SOURCE:
                     (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
 20
             (ix) FEATURE:
                     (A) NAME/KEY: misc feature
                     (B) LOCATION 1...633
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85
25
        AAAAGAAAGA GTATGAAAAG AATGCTGCTG CTTCTCGTTG TATTATTATA TGGAATTGCA
                                                                                                60
        GGCCGATTGG CTGCACAAGA CGTTATCAGA CCATGGTCAT TGCAGGTCGG AGCGGGATAC
                                                                                               120
        TCCGATACGG AGAACATCCC GGGAGGATTC ACCTATGGTT TCTATTTGGG AAAGCGTATG
                                                                                               180
        GGGAGCTTTC TGGAAGTGGG GCTGTCCATG TACAACTCCA CACGTCAAAC AGCCAACAAT
GCAGACTCCT TTGCATCGAA CGAAGGAGAC GGATCTTTTC AGGTAAATAT GTCTTCTCCG
                                                                                               240
30
                                                                                               300
        AATGAGAAGT GGTCATTCTT CGATGCAGGC AGTGCCAACT GCTATATGAT CGTCGTCGGA
GTCAATCCTC TCCATCTGTT TTGGCAGAAT AGCCGGCACA ATTTGTTTCT GGCAGTACAA
GCCGGCCTGT CCAATAAGCA CAATATTCAT TTCATCTATG GAGACAAGGG AGCCAAAGTC
                                                                                               360
                                                                                               420
                                                                                               480
        AGTATCTACA CCAATTCGAA TACCTACATC GGTTACGGAG CACGTGTAGC CTACGAATAT
                                                                                               540
35
        CAAATTCATA AAAACGTGGG GGCGGGTGCC GCTGTAATGT ACGACCACGG CAATAAGATG
                                                                                               600
        CTTACGGCCA TGGCCACGCT CTCCACTCAT TTT
                                                                                               633
        (2) INFORMATION FOR SEQ ID NO:86
40
               (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 2859 base pairs
                     (B) TYPE: nucleic acid
                     (C) STRANDEDNESS: double
45
                     (D) TOPOLOGY: circular
             (ii) MOLECULE TYPE: DNA (genomic)
            (iii) HYPOTHETICAL: MO
50
             (iv) ANTI-SENSE: NO
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
55
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...2859
60
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86
       ATCCGAATGA GAGTATCCGA TCTCTGTTCC AGACTTTCAT GGTTATTACC CGTAATCCTT GTCGGATTGC TCTGTGCTAC TTTGGTCGCT GCGGAACGTC CTATGGCCGG AGCAGTCGGA
       TTGCACCACC GTCGGCATGC TGCGCTGTCT GATTCTACAG CGAAAGACAC GGTGCCTCTC GCAAAACCTA TTCCTGACAG TGCTTTTCGA GATTCCCTTC CTGCCGATTC CACCGGATCG
65
                                                                                               240
       ATGCGGCAAG ATAGCGTGTA TGACGATGAA TTCGAATTGG AAGATATAGT GGAGTACGAA
                                                                                               300
       GCTGCCGATT CCATCGTTTT GCTCGGACAG AATCGTGCCT ATCTTTTCGG CAAGAGCTAT
                                                                                              360
       GTGAGCTATC AAAAGAGTCG CTTGGAGGCA AACTTCATGT ATCTCAATAC CGACAGCAGT
                                                                                               420
       ACGGTTTATA CTCGCTATGT CCTCGATACG GCCGGTTATC CGATGGCCTT TCCTGTTTTC
       AAGGATGGAG AGCAGTCGTT CGAAGCCAAG AACTTTACCT ACAACTTCCG CACGGAGAAG
GGGATTATCA GCGGAGTGAT CACGCAGCAG GGCGAAGGCT ATCTGACTGC CGGTAAGACC
AAGAAGATGC CCGACAATAT CATGTTTATG CAAGGAGGGC GTTATACGAC CTGCGACAAT
                                                                                               480
70
                                                                                              540
                                                                                              600
                                                                                              660
       CACGATCATC CTCACTTCTA TATCAATCTT TCCAAGGCAA AGGTGCATCC GGAGAAAGAC
                                                                                              720
       ATCGTCACAG GTCCGGTCAA TCTGGTTATC GCCGATATGC CGCTGCCGAT AGGTCTTCCT
                                                                                              780
75
       TTCGGCTATT TTCCCTTTTC CAACAAATAC TCTTCCGGTA TATTGATGCC CACGTACGGA
                                                                                              840
```



```
GAGGACAATC GCTATGGATT TTATTTGAGG AATGGTGGAT ATTATTTTGC CTTCAGCGAC
                                                                              900
      TATATCGATT TGGCATTGCG TGGGGAGATC TTTTCCAAAG GGTCATGGGG CATTTCAGCC
                                                                              960
      CAATCGAAAT ATAAGAAGAG GTATAAGTAC AACGGCTCGT TCGAAGCCAA TTATCTGGTA
                                                                            1020
      TCGAAGTCCG GCGACAAATA CGTGCCCGGA GACTACAGCA AGACCACCAG TCTGAATATC
                                                                            1080
 5
      CGATGGACAC ACAGTCAGGA TCCGAAGGCC AATCCTTTGC AAACGTTGTC GGCCAATGTC
      AATTTTGCCA CCGGGAGCTA TTTCCAGAAT TCGCTGAATA CCACCTATGA TGTCAATGCC
      CGTACTGCTA CGACACGAAG TTCGGCCGTG AGCTATTCGC GCAAGTTTCC GGGTACTCCT
                                                                            1260
      TTTTCGATTA CGGGTAGCAT GGATATCAGC CAGAACATGC GCGATACGAC GGTGAGCCTT
                                                                            1320
      ACCTTGCCGA ATCTTTCGAT TAATATGTCC ACGCGTTATC CTTTCAAGCG GAAGACCCGT
                                                                            1380
10
      GTAGGACCGG AGCGATGGTA CGAGAAGTTG AGTGTGGGCT ATTCCGGTCA GCTTCGCAAT
                                                                            1440
      AGTATCTTGA CAAAAGAGAA AGATTTGCTC CAGAGCAATC TCGTGCGCGA TTGGAAGAAT
                                                                            1500
      GGTATGCGTC ATTCCGTACC GATCAGTTTG ACTGTCCCTT TGTTGGATTA TATCAATCTG
                                                                            1560
      ACTATGGGGG TTAACTACAA TGAGTGGTGG TACACGAAAG GCATACGGAA GTCGTGGAAT
                                                                            1620
      GAGGATAAGA AAACATTCCT GCCTTCGGAC ACGACCTATA AATTCCGCAG ACTGTACGAT
                                                                            1680
15
      TACAGTCTGT CGGCAGGCTT ATCTACCACA TTGTACGGTA TGTTCAAGCC TTGGAAACCT
                                                                            1740
      TTTTCCTTCG GAGGCAATCT CATTATGATC CGTCATCGCT TCACGCCCAC TGTCAGTTTC
                                                                            1800
      TCCTATATGC CGGACTTCAC GAAACGCCGA TATGGCTTTT GGGAGCTTCT TGAGCATACG
                                                                            1860
      GATCAGAACG GCAAGCTGCA TACGCTGCTC TACTCTCCTT ATTTCGAGCA GATATTCGGT
                                                                            1920
      GCTCCCTCCA TGGGCAATGC AGGATCTGTC AATTTCTCTT TTGACAACAA CTTAGAGGCC
                                                                            1980
20
      AAGATCAAAT CCAAATCGGA TTCGACAGGG ATCAAGAAGA TCAGCCTGAT AGATCAGTTC
                                                                            2040
      ACATGGTCTA CATCCTATAA TATGTTTGCC GATTCGATCC GATGGAGCAA TATCTCGGCT
                                                                            2100
      TCGCTGGCAC TTCGCCTCTC CAAGAGCTTT ACCTTGCGCT TGTCCGGTCT GTTCGATCCC
                                                                            2160
      TATTTGACGA AGTATTATGA GGGAGAGAT GGGAAGATCA TTCCCTATAA GAGCAACGAC
                                                                            2220
      CTGCGCATTT TTAACGGCAA GGGATTGGCA CGCCTGATCA GTACGGGTAC TTCTTTCAGC
                                                                            2280
25
      TATACGCTCA ACAAAGAGTC GCTCAGCGGA TTGATAGCTC TTTTCAGTGG CAAAAAGGAG
                                                                            2340
      CGGAGAGATG AAAAGAAAAA CACAGGGGCT ACTCCTCATG AAGGAGACGA TGCTGCCGAT
                                                                            2400
      ATACTTGAGG GAGGAAGACC GCAAAATGAA AGTGGGGGGT CGCTCCTCGA GCGCAACCGT
                                                                            2460
      CAGGGCGGAG CAGTGGATCA GGATGGTTAC TTCGCATATT CGATCCCATG GAGCCTGTCC
                                                                            2520
      TTCGACTATA GTTGGAATAT TGCTACCGAC TACAATAGGT ACAATGTCAA TAAGATGGAG CACTACTACC GGGTAACGCA GAATCTGAGC TTTCGCGGCA ATATCCAGCC TACACCGAAC
                                                                            2580
30
                                                                            2640
      TGGAGCTTCG GATTCAATGC GAACTACAAT TTCGACTTGA AGAAAATAAC ATCGCTTACC
                                                                            2700
      TGCAACGTCA CTCGCGACAT GCACTGCTGG GCTATCTCGG CCAGTTTCAT CCCTATAGGA
                                                                            2760
      GCATACAAGT CCTATAATTT CGTCATATCG GTGAAGAGTT CACTCTTGCA GGATCTGAAG
                                                                            2820
      TATCAGCAGA GCAATCGTCC CATCACGAAT ACTTGGTAT
                                                                            2859
35
      (2) INFORMATION FOR SEQ ID NO: 87
           (i) SEQUENCE CHARACTERISTICS:
40
                (A) LENGTH: 3753 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
45
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) AHTI-SENSE: NO
50
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
55
                (A) NAME/KEY: misc_feature
                (B) LOCATION 1...3753
          (xi) SEQUENCE DESCRIPTION: SEO ID NO:87
60
      AAATGGAAAT TGGCATGCGC TTTTGATTGC GCGTGTTGTT TCGATCCATT TGTAGTAACT
      AACGAGGTAA TAATAATGAT GAAACGATAT ACAATAATTC TTGCAGTTTT TCTTTTATTC
      TGCACGGTAT TTACCTTTCA AATAAAAGCT CGCCCTTATG AAAGATTTGC AGATGTAGAG
      AAGCCTTGGA TTCAGAAACA TTCAATGGAT TCTAAATTGG TGCCTGCAAA TAAGGGTAAC
                                                                             240
      TTAATTCAAG CTGAAATTGT ATACCAATCT GTTTCTGAAC ATAGTGACTT AGTTATTTCA
                                                                              300
65
      CCTGTGAACG AAATAAGGCC TGCAAATCGT TTCCCTTCGC ATAGGAAGTC TTTTTTTGCA
                                                                             360
      GAAAATCTAC GGGCATCTCC CCCCGTAGTT CCCGTTGCCG TCGACAAGTA TGCGGTACCG
                                                                             420
      GTTGCCAATC CAATGGATCC TGAAAATCCC AATGCCTGGG ATGTGACGCT AAAAATCACT
                                                                             480
      ACTAAAGCGG TAACAGTACC TGTCGATGTG GTGATGGTTA TCGACCAGTC TTCGTCAATG
                                                                             540
      GGAGGGCAAA ACATTGCCAG ATTAAAGTCT GCCATTGCAT CGGGACAGCG TTTTGTGAAA
                                                                             600
70
      AAAATGTTGC CTAAGGGGAC GGCTACAGAA GGGGTGCGTA TCGCTCTTGT GAGTTATGAC
                                                                             660
      CATGAGCCTC ATCGCTTATC TGATTTTACC AAAGACACTG CTTTTCTCTG TCAAAAAATC
                                                                              720
      CGGGCTTTGA CTCCTATTTG GGGAACACAT ACCCAGGGGG GGCTTAAAAT GGCGAGAAAC
                                                                             780
      840
```

ACGGAGCAGT ATCCTGTTAA AAATGTAACT ACTGCAGACT TCATTGGCAA AACTGGAAAT

GCGAATGATC CCATTGATTT GGTTATACAA GGAGCAATTA ATTTCCCTAC AAATTATGTT

900

960

75



| | TCCAACAATC CATCTACACC TCTTACCCCA AATTATCCAA CTCATTCTTC TAAAGTTGGA | 1020 |
|------------|--|------|
| | CGGAGAAATC TGCCGGAATC CAAATTCGAT TATAGTAATC TGAGTGCAAG GATTACTTTT | 1080 |
| | GATGGTGTTG CTGGCGCATT GGTCTATGAA CCGAGGTTTC CTCATCCCTA TTATTATTAT | 1140 |
| _ | TTCCCTTGTA ACGCTGCTAT CAATGAGGCT CAGTTTGCGA AAAACTCTGG TTATACAATC | 1200 |
| 5 | CATACTATTG GCTATGACCT GGGAGATTTT GCCTTGGCCA ACAATTCGTT GAAACTAACC | 1260 |
| | GCTACAGACG AGAATCACTT CTTTACGGCG ACACCGGCCA ATTTAGCTGC AGCGTTTGAT | 1320 |
| | AATATTGCCC AAACTATTAA TATAGGTATA CAGAGGGGGG AGGTGACGGA CTTTGTAGCT | 1380 |
| | CCTGGTTTCA TCGTTAAAAA TCTGACGCAA TCGGGAGATG TTACTCATTT GCTAAATGTT | 1440 |
| | TCAAATGGAA CGGTGCACTA TGATGTCTCT ACTAAAAAAC TGACATGGAC TACTGGTACT | 1500 |
| 10 | ATCCTGAGCT CATCAGAAGC TACCATAACT TATCGTATTT ATGCCGATTT GGATTATATA | 1560 |
| | CAGAACAATG ATATTCCGGT AAATACTACT TCTGCTATCG GCCCGGATCT TGGTGGATTC | 1620 |
| | GATACCAATA CCGAGGCAAA ATTGACCTAT ACCAATTCCA ATGGCGAACC GAATCAGCAG | 1680 |
| | TTAATTTCC CACGTCCGAC GGTTAAGTTA GGTTATGGTG TTATTAAGCG GCACTATGTA | 1740 |
| | TTGGTAAATA AAGACGGTCA ACCCATACAG GCAAATGGAA CAGTTGTCAG TTCCCTAAGC | 1800 |
| 15 | GAGGCTCATG TTCTACAGTC ACAAGATTTC TTTTTGCCCT CAGGTGGAGG TCATATTGTT | 1860 |
| | CCCAAATGGA TAAAGTTGGA CAAAACGACC GAAGCATTAC AGTACTATTC CGTACCGCCG | 1920 |
| | ACTAACACGG TCATCACTAC TGCCGATGGT AAACGTTATC GTTTTGTCGA AGTCCCAGGC | |
| | TCCACGCCGA ATCCGGGCCA AATCGGTATC AGTTGGAAAA AACCGGCAGG AAACGCTTAC | 1980 |
| | TECHOCOGO A ACCOGORAL MALEGGIAIC AGTIGGAAAA AACCGCCAGG AAACGCTTAC | 2040 |
| 20 | TTCGCTTACA AGCTCCTCAA TTATTGGATG GGAGGAACAA CAGACCAACA GAGTGAATGG | 2100 |
| 20 | GATGTGACGT CCAATTGGAC AGGAGCCCAA GTACCGCTCA CAGGAGAAGA TGTAGAGTTT | 2160 |
| | GCAACGACAG AAAATTTCGG TTCTCCGGCG GTAGCCGATT TGCATGTCCC GACAACCAAC | 2220 |
| | CCCAAAATTA TCGGTAACCT TATCAATAAT TCCGACAAGG ATTTAGTTGT TACCACAAGC | 2280 |
| | AGTCAATTGA CGATCAACGG CGTGGTTGAG GATAACAATC CGAATGTCGG TACGATCGTC | 2340 |
| 0- | GTGAAGTCGT CGAAAGACAA TCCTACGGGG ACATTGCTTT TTGCCAATCC GGGCTATAAT | 2400 |
| 25 | CAAAATGTAG GGGGGACCGT CGAGTTTTAC AATCAGGGAT ATGATTGTGC CGATTGTGGT | 2460 |
| | ATGTATCGCA GGAGCTGGCA GTATTTCGGT ATCCCTGTCA ATGAATCAGG TTTTCCAATT | 2520 |
| | AATGATGTGG GCGGAAACGA GACCGTCAAC CAATGGGTTG AGCCTTTCAA TGGCGATAAG | 2580 |
| | TGGCGGCCAG CACCTTATGC ACCTGATACA GAGCTTCAAA AATTCAAGGG CTACCAGATC | 2640 |
| | ACGAATGACG TGCAGGCACA GCCTACGGGA GTTTACAGCT TCAAGGGTAT GATTTGTGTG | 2700 |
| 30 | TGCGATGCCT TCCTGAATCT GACACGCACG TCCGGTGTCA ACTACTCGGG CGCCAACTTG | 2760 |
| | ATCGGCAACT CATACACTGG AGCCATCGAC ATCAAGCAGG GTATTGTCTT CCCGCCGGAA | 2820 |
| | GTCGAGCAGA CGGTGTATCT GTTCAACACG GGAACACGCG ACCAGTGGCG TAAGCTTAAT | |
| | GGAAGCACGG TTTCAGGCTA TCGAGCCGGT CAGTACCTCT CTGTACCTAA GAATACAGCG | 2880 |
| | GGTCAGGACA ATCTTCCGGA TCGTATTCCA TCGATGCATT CCTTCTTGGT GAAGATGCAG | 2940 |
| 35 | AACGGAGCGT CTTGTACGTT GCANATCTTG TACGATAAGC TGCTCAAGAA CACGACTGTA | 3000 |
| 00 | AACOAGCGT CITCHACGT GCANAICTG TACGATAAGC TGCTCAAGAA CACGACTGTA | 3060 |
| | AACAACGGTA ATGGTACGCA GATCACATGG CGATCCGGCA ACTCCGGATC GGCGAATATG | 3120 |
| | CCGTCACTTG TGATGGATGT TCTTGGTAAC GAGTCGGCCG ACCGTTTGTG GATCTTTACC | 3180 |
| | GATGGGGGTC TTTCTTTCGG ATTCGACAAC GGCTGGGATG GTCGCAAGCT GACTGAAAAA | 3240 |
| 40 | GGTTTGTCAC AACTTTATGC GATGTCTGAC ATCGGTAATG ATAAATTCCA GGTTGCAGGG | 3300 |
| 4 0 | GTTCCGGAGT TGAATAACCT GCTGATCGGC TTCGATGCGG ATAAGGATGG TCAATACACG | 3360 |
| | TTGGAGTTTG CTCTTTCGGA TCATTTTGCG AAAGGGGCTG TTTACCTGCA CGATCTTCAG | 3420 |
| | TCAGGAGCCA AACACCGTAT TACGAATTCT ACGTCGTATT CATTCGATGC CAAGCGGGGA | 3480 |
| | GATTCCGGGG CTCGTTTCCG CTTGTCATAT GGATGTGATG AGAACGTAGA TGATTCGCAT | 3540 |
| | GTCGTGAGTA CAAATGGCCG TGAAATTATA ATTCTGAATC AAGATGCTCT TGACTGCACT | 3600 |
| 45 | GTAACCTTAT TCACAATAGA AGGTAAGCTT CTTCGCCGCT TGAAAGTATT AGCTGGTCAT | 3660 |
| | AGAGAAGTCA TGAAAGTGCA GACCGGAGGG GCCTATATTG TGCATCTTCA AAATGCTTTC | 3720 |
| | ACTAATGATG TGCATAAGGT GCTTGTTGAG TAT | 3753 |
| | The second secon | 3733 |
| | | |
| 50 | (2) INFORMATION FOR SEO ID NO:88 | |
| | , and a state of the state of t | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 1278 base pairs | |
| | | |
| 55 | (B) TYPE: nucleic acid | |
| 00 | (C) STRANDEDNESS: double | |
| | (D) TOPOLOGY: circular | |
| | | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| CO | | |
| 60 | (iii) HYPOTHETICAL: NO | |
| | | |
| | (iv) ANTI-SENSE: NO | |
| | ()) 000000000000000000000000000000000 | |
| 65 | (vi) ORIGINAL SOURCE: | |
| oo | (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| | | |
| | (ix) FEATURE: | |
| | (A) NAJE/KEY: misc_feature | |
| 5 0 | (B) LOCATION 11278 | |
| 70 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88 | |
| | | |
| | ACTATGGAAG TGAAGAAAAA CACAGTGGTG CTACGCCTTC TGATTTGGTT CGTGGCCATT | 60 |
| | CTTCTCTTCC ACTCCTCACG GCTGTGGGGA CAGGAAGGGG AGGGGAGTGC CCGATACAGA | 120 |
| <i>7</i> 5 | TTCAAAGGAT TCGTGGATAC CTACCATGCC GTACGCAGCT CTTCTCCTTT TGATTTCATG | 180 |
| | | |

```
AGCTCGCGTA CGAGAGTGAG AGGTGAGCTG GAGAGGTCGT TCGGTAATTC GAAAGTAGCC
                                                                                    240
        GTATCGGTCA ATGCCACCTA CAATGCTCTA CTGAAAGACG AGACCGGCTT ACGTTTACGT
                                                                                    300
        GAAGCCTTCT TCGAGCATCA GGAAGAGCAT TGGGGGTTGC GCCTCGGACG ACAGATTGTC
                                                                                    360
        ATTTGGGGGG CTGCCGACGG TGTGCGCATC ACGGATCTGA TCTCCCCGAT GGATATGACC
                                                                                    420
  5
        GAGTTTCTGG CACAGGATTA CGATGATATT CGTATGCCGG TCAATGCATT GCGTTTCTCT
                                                                                    480
        GTCTTCAACG AATCGATGAA AGTGGAAGTC GTGGTACTGC CTGTATTCGA GGGGTACCGT
                                                                                    540
        CTGCCTGTGG ATCCTCGCAA TCCTTGGAAT ATCTTCTCCC TTTCGCCCAT TGCTCAGGGG
                                                                                    600
        ATGAATATCG TCTGGAAAGA AGAAGCCGGC AAACCGGCCT TCAAGGTTGC CAATATCGAG
                                                                                    660
        720
 10
       ACATGGAACA AGATGCCCGT CATCGAAGTA CAGGGCATTG TGCCGACGGA AATCATCGTT
                                                                                   780
        AGCCCTCGCT ATTATCGTAT GGGATTTGTC GGCGGCGACC TCTCCGTACC CGTCGGACAG
                                                                                   840
       TTTGTTTTCA GGGGAGAGGC TGCGTTCAAT ATCGACAAAC ACTTCACCTA TAAGAGTCAT
                                                                                    900
       GCCGAGCAAG AGGGTTTCCA AACAATCAAT TGGTTGGCCG GAGCCGATTG GTATGCTCCC
                                                                                   960
       GGTGAATGGA TGATCTCAGG ACAATTCTCA ATGGAAAGCA TATTCAGGTA TAGGGATTTC
                                                                                  1020
       ATCTCCCAAA GACAACATTC TACCCTGATT ACTCTCAATG TTTCCAAGAA ATTCTTCGGC
15
                                                                                  1080
       AGTACACTCC AACTTTCGGA CTTCACCTAC TACGACCTTA CGGGCAAAGG ATGGTTCAGT
CGCTTTGCAG CTGACTATGC CTTGAACGAT CAGATACATC TGATGGCCGG ATATGACTGG
                                                                                  1140
                                                                                  1200
       TTCAGTAGTA AGGGCAGCGG TATATTCGAT CGCTACAAAG ACAATTCCGA ACTCTGGTTC
                                                                                  1260
       AAAGCCCGCT ACAGCTTC
                                                                                  1278
20
       (2) INFORMATION FOR SEQ ID NO:89
             (i) SEQUENCE CHARACTERISTICS:
25
                  (A) LENGTH: 1392 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
30
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
35
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
40
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1392
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89
45
       TACGGCAAAA GAAGAAAGCT CGGAACCTCC GTCCGACCTT CCGTTCTAAC CCAAATCAGA
       TTTATTTTGG ACTTACACCT TATTACTGAT TTTTTTTGAAG GGCTTCGGGT CAATCCTATC
GGTGCAGCAG CCATAGTGGC TTTCATTATC GACCTGCTTC TTCTTTGCTG TTCGGCTTTT
                                                                                   120
                                                                                   180
       ATGTCCTCCT GTGAGGTGGC TTATTTTCA CTAAAGCCGA TCGATCTGCA GAACATCCGC
                                                                                   240
       GAACGGAATC ACTCTTCCGA CATCGCGCTT TCCAATTTAT TAGACAATTC GAATCAGCTA
                                                                                   300
50
       TTAGCTACTA TTCTGATCGG GAATAATGTG ATTAATGTAG CCATCGTTAT CCTTTCCAAT
                                                                                   360
       TATGCCATCG AGCAGACATT CGTTTCTCT TCTCCGATCA TTGGATTTCT GATCCAGACG ATACTCCTGA CCACTGTTCT TTTGCTGTTC GGAGAGATTC TGCCGAAAGT GTATGCGCGG
                                                                                   420
                                                                                   480
       AAGAATCCGC TGCAATACTC GCGCTTTTCT GCTGCAGCTA TGTCCGTTAT CTATAAGATA
                                                                                   540
       TTGTCACCGT TTTCAAAATT GCTGGTCAAA AGTACCGGCA TCGTTACCAG AGGTATCAGC
                                                                                   600
55
       AAGAAGAAAT ACGATATGTC CGTGGATGAG CTCTCGAAAG CGGTAGCCCT CACCACTACG
                                                                                   660
       GAGGGAGAGC CGGAGGAGAA AGAAATGATT AACGAAATCA TCAAATTCTA TAATAAGACA
                                                                                   720
       GCCTGCGAAA TCATGGTTCC GCGTATCGAT ATTGTGGATG TGGATCTGAG CTGGCCATTT
CGTAAGATGC TTGACTTCGT TGTTTCGTCG GGTTATTCCA GACTTCCCGT TTCAGAGGGG
                                                                                   780
                                                                                   840
       TCAGAAGACA ATATCAAAGG GGTGATTTAC ATCAAAGATC TAATCCCACA CATGGATAAA
                                                                                   900
60
       GGCGATGAAT TCGACTGGCA TCCTCTGATT CGTAAAGCAT ATTTTGTCCC CGAAAACAAG
                                                                                   960
       CGCATAGATG ATTTGCTCGA GGAGTTCAGA GCCAATAAGG TGCATGTCTC CATCGTTGTG
                                                                                 1020
       GATGAGTTCG GTGGCACTTG CGGACTGATC ACAATGGAGG ACATATTGGA AGAGATCGTC
                                                                                 1080
       GGCGAGATTA CGGACGAGTA CGATGAGGAA GAACTCCCCT TTAAGGTTTT GGGGGATGGC
                                                                                 1140
       AGTTATCTTT TCGAAGGAAA AACGTCTCTC TCCGATGTTC GACACTATCT TGACCTTCCG
                                                                                 1200
65
       GAAAATGCTT TCGGTGAATT GGGGGACGAG GTAGATACGC TAAGTGGGCT CTTCTTGGAA
                                                                                 1260
       ATCAAGCAGG AACTCCCCCA TGTGGGCGAT ACAGCAGTGT ACGAGCCATT CCGCTTTCAA
                                                                                 1320
       GTGACCCAAA TGGACAAGCG CCGAATCATC GAAATCAAGA TTTTCCCTTT CGAGCGCACT
                                                                                 1380
      TGGGAGGTCG AA
70
       (2) INFORMATION FOR SEO ID NO: 90
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 798 base pairs
75
                 (B) TYPE: nucleic acid
```

```
(C) STRANDEDHESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
 5
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
10
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
15
                 (B) LOCATION 1...798
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90
      ATAATTTATA GATCAACAAT GAAACTATTA CTTTATCTCC TATTGGTCTT GTCGACTCTA
20
      TCCCCGATGT ATTCGCAAAT GCTCTTCTCA GAGAATCTCA CAATGAATAT AGACAGCACC
                                                                                   120
      AAAACCATAC AAGGAACGAT ATTGCCCGTA CTGGATTTCA AAACCGAAAA GGAAAATGTG
                                                                                   180
      TTCACCTTCA AAAATACTGC CAATCTCAAT CTGCTGATAA AGCACGGTCA AGTAATCAAC
TTAATTAATA AGCTTGAGTT TTCTACCTAT GGCAATAAAG TAACCGTAAG TGGAGGATAT
                                                                                   240
                                                                                   300
      GTACACACCG AATACCGCTA TTTGTTGCAT CATGTTTTTG AGGTTTATCC TTATGTCGAG
                                                                                   360
25
      TCGCAATGGG CAGAAAGTAG AGGAATGAAA TATAAGGTTT CTACGGGATT ACAGTCGCGT
                                                                                   420
      TATCGGCTGG TAAATAGTGA TAACTGTCTC ATGTTTGCAA CATTGGGGGT ATTTTTCGAA
                                                                                   480
      TTCGAAAAGT GGGAACAGCC AGCCACTAGC CTCTTTGCAG GAACGTATGC ATACAGCCGA
                                                                                   540
      AGTATCAAAA GCCACCTGTC TATCAGTTTC AGACATCGGT TGGGTGAACA TTGGGAATTT
                                                                                   600
      ACAACTACGG CTATTCACCA GGGAAAGCCT GACAGTTATT TTAAGAAGGC ACGTTTTGGA
                                                                                   660
30
      GGAGCTATCG ACCTCAAATA CCATATCACA CCTACGATAG GAATACGCGG GGCCTATCGG
                                                                                   720
      ATCATCTACG ATACTGCCCC TATTGTACCT GTGCGGAAAG ATTACAACAC CGTTGATGTT
                                                                                   780
      GGTATCGATA TTTCGTTT
                                                                                   798
35
       (2) INFORMATION FOR SEQ ID NO:91
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2721 base pairs
                  (B) TYPE: nucleic acid
40
                  (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
45
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
50
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...2721
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91
      ACTATATGCG TTGCCTTCTT ATCCGCTCCC GTTGCGGCAT TGCGCGAGTC GCCGCCAATG
                                                                                    60
      GGGGCGGAGC GGAAGACTCC CTCTCTCTA CCCCTTCTTT TCGAGCATTC TGATAAAGGA
                                                                                   120
60
      AAGGGGTTCG ACTACAGACT ATTCACAAGC AATAAACTTA AAGTGTTTGC CACAGGTAAC
                                                                                   180
      AGCCGATACA TACACAATAA ACCAACAATC ATCCAAGCAA TGAAACGAAT CGTTTTATCA
TCTTTCCTGT TCGTTCTGTC CATACTTTCT TTGATGGCAC AGAACAATAC CCTCGATGTA
                                                                                   240
                                                                                   300
      CACATATCCG GTACGATCAA GGATGCCTCC TCCGGCGAAC CAGTGCCCTA TGCCACTGTA AGCATCCGGC TGACAGGAGC AGATACCACA CAGGTGTTCC GACAAGTGAC TGACGGCAAC
                                                                                   360
                                                                                   420
65
      GGCTACTTCG TCATAGGCCT GCCGGCAGCT CCCTCCTATC ACCTGACAGC TTCGTTCGTA
                                                                                   480
      GGTATGAAAA CCCATACCAT GCAGATTAGT CGGGGAAATG GACAGCACGA CATCAAATCC
                                                                                   540
      ATCGACATTT CTCTCGAATC CGAGGACAAA CAACTCTCCA CCGTCACCGT ATCGGCAGCA
                                                                                   600
      CGACCACTGG TGAAGATGGA GATAGACCGC CTGTCCTATA ATATGAAAGA TGACCCCGCA
                                                                                   660
      GCCAAGACGA ACAACCTGCT CGAAATGCTG CGCAACGTTC CTTTGGTAAC GGTGGATGGT
                                                                                   720
70
      CAGGGCAATA TCCAGGTGAA AGGATCTTCC AACTTCAAAA TCCACCTCAA TGGCAGGCCC
                                                                                   780
      TCGACCATGG TGAGCAGCAA CCCGAAGGAG GTCTTTCGCT CCATTCCTGC CCATACGATC
                                                                                   840
      AAACGGGTGG AGGTCATCAC CGATCCGGGT GTAAAGTACG ATGCGGAAGG CACAAGTGCC
                                                                                   900
      ATCCTGGACA TCGTCACGGA AGAAGGTAAG AAGCTGGAAG GATATTCAGG TTCCATCACG
                                                                                   960
      GCCAGTGTCA GCAACAATCC CACAGCCAAC GGTAGTATCT TTCTGACGGC AAAGTCCGGC
                                                                                  1020
75
      AAAGTCGGGC TGACTACCAA CTATAACTAC TACGGTGGCA AAAACAAGGG CTCTCGCTAC
                                                                                  1080
```

WO 99/29870

```
TTTACCGAAC GTACTACATC CATGCTCCAA ACGATAGAAG AAGGCAAAGG GCAAGAAACC TTTGGCGGAC ACTTCGGCAA TGCCCTCCTC TCATTCGAGA TAGATTCGCT CAATCTCTTT
                                                                                     1140
                                                                                     1200
       ACGGTGGGCG GCAATGTACG CCTTTGGGAG ATGACCACCG ACCGGAACAG CGTAGAAAAA
                                                                                     1260
       AGCTTTGCCG GCAGCAACCT CATGTCCTAC ATAGACAGAA AACTCAAAAC ACAGATGGAT
                                                                                      1320
 5
       CTCACCGTTT CCTACCGCTT CACTCACAAT CCTAATAATA GCGAGACCTT CATTGACCAA
                                                                                     1440
       TGGAAGCGCG ATCCGCTCAA CACAGCTAAT ACGATCCAGT ACGCCGGCCA GCACTCCAAA
                                                                                     1500
       TCCGATGCGG GCATGGACGA ACATACGGCA CAAGTGGACT ATACACGTCC CTTAGGACAA
                                                                                     1560
       GCACATTCTT TGGAAGCAGG GCTGAAGTAC ATCTATCGTC ATGCCACGAG CGATCCTCTC
                                                                                     1620
10
       TATGAGATAC GACCATCCGA AGATGCTCCG TGGCAGCCCG GCTCTCTATA TGCACAGAAT
                                                                                     1680
       CCGTCGAACG GAAAGTTCCG CCACGATCAA TACATCGGAG CAGCCTATGC CGGCTACAAC
                                                                                     1740
       TATCGTAAGG ATCAGTATTC TTTGCAAACC GGCCTCCGAG TGGAAAGCAG CAGGCTGAAA
                                                                                     1800
       GCACTCTTTC CCGAAAACGC AGCAGCAGAT TTCTCCCACA ACTCGTTCGA CTGGGTGCCA
                                                                                     1860
       CAGCTCACGC TCGGCTATAC CCCCTCGCC ATGAAGCAGC TTAAGCTGGC CTATAACTTC CGAATCCAAC GTCCTGCAAT CGGCCAACTG AATCCCTACC GGCTACAGAC CAACGATTAT
                                                                                     1920
15
                                                                                     1980
       CAAGTACAGT ATGGTAATCC CGACCTAAAG TCGGAGAAGC GTCACCACGT CGGTCTCTCC
                                                                                     2040
       TATAATCAAT ACGGAGCCAA GGTCATGCTT ACAGCATCGC TCGACTACGA CTTCTGCAAC
                                                                                     2100
       AACGCCATCC AGAATTACAC CTTCTCCGAC CCGGCCAATC CCAATCTGTT CCACCAGACC
                                                                                     2160
       TATGGCAATA TCGGACGAGA GCATTCTTTC AGCTTGAATA CCTATGCCAT GTACACGCCG
                                                                                     2220
20
       GCCGTATGGG TCAGGATTAT GCTCAACGGA AATATCGATC GCACATTCCA AAAGAGCGAA
                                                                                     2280
       GCACTCGGCA TTGATGTCAA TTCATGGTCC GGCATGGTAT ACTCAGGCCT GATGTTCACC
                                                                                     2340
       CTGCCGAAGG ATTGGACTGT GAATCTCTTC GGAGGTTATT ATCATGGGGG AAGAAGCTAC
                                                                                     2400
       CAGACGAAGT ATGATGGCAA TGTATTCAAC AATATCGGTA TAGCCAAACA GCTTTTCGAC
                                                                                     2460
       AAAAAATTGA GAGTCTCGCT GAGCGCAAAC AACATTCATG CGAAGTATTC GACATGGAAG
                                                                                     2520
25
       AGCCGGACCA TCGGCAATGG ATTTACTATT TATTCGGAAA ATGCCGGTAT ACAACGGAGT
                                                                                     2580
       GTTTCCCTCA GCCTCACCTA CAGCTTCGGT AAGATGAATA CACAAGTGCG CAAGGTAGAG
                                                                                     2640
       CGTACGATCG TCAACGACGA CCTCAAGCAA ACCTCATCCC AAGGACAGCA GGGTGGCGGA
                                                                                     2700
       CAAGGAAATC CTACCGGCAA T
                                                                                     2721
30
       (2) INFORMATION FOR SEQ ID NO:92
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1350 base pairs
35
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                   (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
40
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
45
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
50
                   (B) LOCATION 1...1350
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92
       TGGTGCCAAT CCGACCCAAT ACCTCATTGC CATGCAGTAT ATCGAAACGC TCAAGGACAT
55
       CAACAAGGGC GATCAGACCA AGACCGTTTA CCTCCCGTTC GAGGCTACCG GTATGCTCGG
                                                                                      120
       TGCCTTGGGC GGTATGAAGG ATTTGGTGAA AGGATGAGAC TCTCTGCCAT TCTTATCGCT
                                                                                      180
       TTGATTGTGA TGCTGCCTGC TGTGCTTAGC GGGCAGCATT ATTATTCCAT GGCGGGAGAG
                                                                                      240
       CGACTGGAGA CGGACAGCAT TCGTCCGAAC GAACTCTCGG CATCGATCCG AAGTGCGCTT
                                                                                      300
       TTCTTTCGGA ACAATGAATA CAATGCACGT TCGGTCAAAG GTTATACGTT GCCGGGTGCA
                                                                                      360
60
       CGGGTTTCCG CTTTTGCCTC TTACTCGCTG CCGGCAGCAC ATGGTGTGAA GCTTTCGCTC
                                                                                      420
       GGAGTATCTA CCCTGAACTA CTGGGGGGCA AGTCGCTATC CGGCCGGTAT CGCTTATTCC
                                                                                      480
       GATTTACCTT ATTGGACGGA CTATAACGAC TATGTACGCT TGCGTATCCT GCCTTATGTA CAGGCCATGC TGAAGCCGAC GGCCACGACT GCTCTCATGC TGGGCAATAT AGCCGGTGGT
                                                                                      540
                                                                                      600
      ACGGCTCACG GACTGATCGA ACCGATCTAC AATCCTGAGT TGGATTTGAC GGCTGATCCT
GAAGCCGGTG TGCAATTTCG GGGTGATTGG ACACGTTTCC GAATGGATGT TTGGGTCAAT
                                                                                      660
65
                                                                                      720
       TGGATGAGCA TGATTTTCAA AAATGACAAT CATCAGGAGT CGTTTGTCTT TGGCTTGTCC ACTACTTCGA AATTGTTATC GGGTGAAGGC AAATGGCGAC TCGAACTGCC CTTGCAGGCT
                                                                                      780
                                                                                      840
      ATTGCCACGC ATCGCGGCGG GGAATACAAC TGGGCGCAGC AGGATACCGT GCATACATGG
GTCAATGGAG CTGTCGGACT TAAGCTTTCG TATCGCCCTC GTACCGACAA ACCCATGCAG
ATTTGGGGAT CTGCTTATGG TGTGGCAGCC TTGTCAAGCG GAGGATACTT CCCTTACGAA
                                                                                      900
                                                                                      960
70
                                                                                     1020
       AGAGGGTGGG GCGGTTATCT TTCTCTCGGA ATGGACTTGG AGCACTTCGC TTTTCGTACC
                                                                                     1080
       GACTATTGGT ACGGCAGGCA TTACGTTTCT CCCTTTGCTG CACCTTTCGC CAATTCCCTG
                                                                                     1140
       ACGTATGACA AACAGCCTCT TACGAACGGT TGGGGCGATT ATATTCGTCT CTATGCCGAC
                                                                                     1200
       TATTCGTGGC GGATGGCACG AAGTGTTTCG TTGGCGGCTG TTGCTCGGGT ATGGTTCCAG
                                                                                     1260
75
       CCTTCGGATC GTTTTGCGAT GAGCCACGCC TTGGAACTGA CGATGCGTAT CGATCCCAAA
                                                                                     1320
```

| | TTCCCAATAG CTTTTCTGAA AGGCAATCAT | | | |
|------|--|--|--|--|
| _ | (2) INFORMATION FOR SEQ ID NO:93 | | | |
| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1341 base pairs (B) TYPE: nucleic acid | | | |
| 10 | (C) STRANDEDNESS: double(D) TOPOLOGY: circular | | | |
| | (ii) MOLECULE TYPE: DNA (genomic) | | | |
| 15 | (iii) HYPOTHETICAL: NO | | | |
| 10 | (iv) ANTI-SENSE: NO | | | |
| | (vi) ORIGINAL SOURCE: | | | |
| 20 | (A) ORGANISH: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAHE/KEY: misc_feature (B) LOCATION 11341 | | | |
| 25 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93 | | | |
| 30 | CCGTTTTCAC CGCCATGGGT ATGTGCTGGA GCGGCTTCCT CAGTACGCAC ACGGCCATGC TCGACAGTCT CAAGTACCGC CGCCTCATCA CACAGGCGAT CGGAGCACAT GCGATCGGCG GACTCGTCGC CGGCATATTC GCCCATTGGC TCTTTGTCCT CATCTCACTC ATTTGATTCG ATACCCGACC AACGTATGAA CAAATCGCTA TTATCATTGG CATGCCTCAT CCTGTGCGGT ATGCCGGCCA TCGCCCAACA GACAGGACCG GCCGAACGCA GCGGCGAGCC TTCTCTGGCC GAACGTGTAT TCGGTCTGGA GCAGAAGCAG AAAAAGCTGA AGGTGTACTT AGGCATACAG | 60 120 180 240 300 360 | | |
| 35 | TCGTTCTACG ACCAGCCGCT TGTCGATGAC GAATCCCATA TCGGACACTT CAAGGTACAG GAGCTGCGGA TGTCTGCTCA TGGCGAACTG AACCGCCACC TCAGCTTCGA CTGGCGACAA CGTCTCAACC GTGCCGCCGA CGGCACTTCG TTTGCCGACA ATCTCTCCAA TGCCATCGAC ATCGCAGGTG TGGACTGGCA CCCGAACGAC AAGGTGTCTT TCTTCTTCTGG ACGTCAGTAC GCGCGTTTCG GAGGGATAGA ATACGACATG AACCCCGTAG AGATCTACCA GTACAGCGAC | 420 480 540 600 660 | | |
| 40 | CTTGTGGATT ACATGACCTG CTATACTTCG GGCGTGAACT TCGCATGGAA CTTCCACCCC GAACAGCAGC TGCAGCTACA GGTACTCAAT GCTTACAACA ACCGCTTCGC CGACCGCTAC CACGTGACAC CCGATGTCGC TACCGCCACG AGCTACCCG TCCTCTCTC GGCACAGTGG AACGGTACCC TCCTCGGAGG AGCACTGCAT ATGCGTTACC CCGTGTCGAT GGCTCATCAG GCCCAAGAGC GTAATATGTG GTACTTCACT GCGGCAAACCG | 720 780 840 900 960 | | |
| 45 | ATCAACGGAT ACCTCGACCT CACCTACTCG ATCGAGGGAT TGGACGACAA AGGCATTATG ACTGCTCGCT ACGGCAAGG CAAGACCCTC ACGGACGTCA AGTACTATGC TCTGGTATCG AAGTGGAACT TCCGCAATTAT CGACCAGGTC AATCTCTTCC TCAAAGGCAT GTACGAGAAC GGCTATGCGC CGCCAATAA CGGCGAGAGC AGCCACACGC GCCACTCCTA CGGCTATATG GGAGGGGGGG AATATTACCC TACGGAGACC AACTTCCGTC TGTTCGTCAC CTACATAGGA CGGCATTACC GGTACAGTGC GACCGAGACG GAAAGCACCA ATGCTCTTCG CGCCGGTCTG ATCTATCAGA TACCTTTCTT A | 1020 1080 1140 1200 1260 1320 1341 | | |
| 50 | | | | |
| | (2) INFORMATION FOR SEQ ID NO:94 | | | |
| 55 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 681 base pairs (B) TYPE: nucleic acid (C) STRANDEDMESS: double (D) TOPOLOGY: circular | | | |
| 60 | (ii) NOLECULE TYPE: DNA (genomic) | | | |
| | (iii) HYPOTHETICAL: NO | | | |
| C.F. | (iv) ANTI-SEHSE: NO | | | |
| 65 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | | | |
| 70 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1681 | | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94 | | | |
| 75 | TATAAGCAAG CTATAATGAA ACATTTGTTT AAGTCGACAT TAGTACTTCT TTGTGCTCTT | 60 | | |

| 5 | TCTTTTTCCG GTACCTATAC CTTTGCACAA GAAAATAATA CAGAAAAGTC ACGATTTGAT TTTTCTGTTA GGCTGGACA GGGATATATT GCAGGTTCAA CTACCAACCT GATGTATGGG TATACATCTG CTAACGATAG ACTTTTGTCT GGTGCAATTT ATCTGGGCTT GACACCAAGT AAGAAAGAAAA ATGCAACCG CGTAGCATTT CGTTCTTAT CHCCCTCTCC GGGTTATTAT GTCGATATAT CCGGCAAAGA AAATACCTTG AATTATGCGT TTTACGTTGT CGGAGCATAT AATAGAATAG CCATTCCTAT ACGCCCTATC AAAAAATTTTA ATTTCATCTT CTCTACAGAA GTCGGAAATG CTTGGATGAG CAAATTTACA ATTCTACTTC GCAGACTTTGG | 120 180 240 300 360 420 480 |
|------------|--|---|
| 10 | GATAAGCAGC GCAAGTCGAG GTCGGGACTG GATTTTGGTC TCGGGATGCA TCTGCAATHC CACATTAATA AGACCGTTTA CTTTATGGCA GGAACCGATC TTACGTCTTG CATGTTCGGA AAAAGGATCA ATGACTACCA GCAAAAGGAT CGAACCTTCA TTGCACTTAT CGACAACAGT ATTGGCATAG GATTAAACCT C | 540 600 660 681 |
| 15 | (2) INFORMATION FOR SEQ ID NO:95 | |
| 20 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1218 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 25 | (iii) HYPOTHETICAL: HO | |
| 25 | (iv) ANTI-SENSE: NO | |
| 30 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 30 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11218</pre> | |
| 35 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95 | |
| 40 | ATATTCATAG ACCCCGATAA GAATACAAAA CAAAACGAAC GAAATATGAT TATCAAGAAA ATACCTGAAAA ATAAATTGGC CCCCTTGGCC ATACTGTTCC TTTTTGCTCC AAAGGCTATG AAGGCTCAGG AGCAACTGAA TGTGGTACAC ACCTCTGTGC CATCGCTGAA TATCAGTCCG GATGCACGTG CGGCCGGTAT GGGGGATATA GGTGTGGCAA CGACGCCGGA TGCGTATTCA CAGTATTGGA ATCCGAGTAA ATATGCTTTC ATGGATACGA AAGCCGGTAT TAGCTTCTCA | 60 120 180 240 300 |
| 4 5 | TATACACCCT GGCTGTCCAA GCTGGTCAAT GATATTGCCC TGATGCAGAT GACCGGTTTC TACAAATTGG GAACAGACGA GAATCAGGCT ATTAGTGCTT CTCTGCGTTA TTTCACATTA GGAAAGTTGG AGACTTTCGA CGAATTGGGC GAATCCATGG GAGAGGCCCA TCCCAATGAA TTTGCTGTCG ATTTGGGCTA TAGCCGCCAG TTGTCGGAGA ACTTCTCCAT GGCTGTTGCA CTGCGTTACA TCCGCTCAGA CCAAAGCACT CACAACACCG GAGAGAATCA GGCCGGAAAT GCCTTTGCGG CGGATATAGC CGGTTATTG CAGAAGTATG TGCTACTGGG TAATGCGGAG AGCTTGTGGT CGTTGGGTTT CAACGTAAAG AATATCGGAA CGAAGATCCC CTATGACGGA | 360 420 480 540 600 660 720 |
| 50 | GGTGTCACGA GTTTTTTCAT CCCTACTTCG TTGAATCTCG GGACGGGGCT GTTGTATCCG ATCGATGACT ATAACAGCAT CAATTTCAAC CTTGAACTTA GCAAGCTGCT TGTACCCACT CCTCCTATCA TGGATCAAAC CGATCAGGCC GGGTATGAGG CTGCACTCAA GAAATATCAG GAAACTTCTT CGATCAGCGG TATATTCTCT TCTTTCGGTG ATGCGCCGGG AGGACTCAAG GAAGAATTCC GTGAGATTAC ATGGGGACTT GGGGCTGAAT ATAGCTATGA CGATAAATTT | 780 840 900 960 1020 |
| 55 | TTTGTTCGTG CCGGATATTC ATACCTGCAC CCCACCAAAG GCAATTTGCA GTACTTCACG GCCGGTGCCG GCTTCAAAAT GAACATATTC CGTATCGATG CTTCCTACCT GTTGTCTACG ATCCAGAGTA ATCCGTTGGA TCAGACTCTG CGGTTTACGC TTGCTTTCGA TATGGATGGA TTGCGCAATT TGTTCCAC | 1080 1140 1200 1218 |
| 60 | (2) INFORMATION FOR SEQ ID NO:96 | |
| 65 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1356 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 70 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SEMSE: NO | |
| 75 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROHONAS GINGIVALIS | |

```
(ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...1356
  5
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96
        CTAATAATCG AAAAGGAAAT GAAAACAACA GTTCAACAAA TTATTCTGTG CCTGGCTTTA
ATGATGTCAG GTGTATTGGG CGGAAACGCA CAGAGCTTTT GGGAAGAAAT AGCTCCTCCT
                                                                                           120
10
        TTTATCAGTA ATGAGCCTAA CGTCAAGTAT ATAATTCCCA ATATGGGGAT TGATTCAAAG
                                                                                           180
        GGAACAATCT ATGTAACCGT GACAAAAAGG ATTCAGCAGG GAGCAAATTA TACTTCTGAG
        CAATTGGGTA TGTACTATCG ACCATTAGGT GATAATGAAC AGTGGTGGAA ACATGATCCG
        TATTTTGATG ACAAGATAGT TGCGGATATT CAGACAGATG CATATGGCAG AGTTTATGTA
                                                                                           360
        TGTACGACTT CTTCTCGAGA TCAAGAGTAT CAACTTTATA TAAACGAGCA GAACGAATGG
AGGTGTATAT TCAAAACTTC TGTGTCTACA TATGAGCATG GTATGGCTGT TTTTCGCTCT
                                                                                           420
15
                                                                                           480
        TCGACAGGGG TGACTTATAT AGGTACCAGG CATCACATCT TCGCATCAGG TGTAAATGAT
                                                                                           540
        TTCGAGTTCA ACACTATCTA TGAAGACTCT ACACCTATGA GCTGTCGCTT TGCAGAGGCT ACGAATAGTG GCACCATCTA TCTGGCATTG ATGCATGAAA CCACAATGTC TACGACTATC
                                                                                           600
                                                                                           660
        CTTACTTATC AAAACGGTGA GTTCGTCGAT ATCTCGGAAA GTGAATTGAG TAACTCGATT
ATTGCATCCA TGTGCTCTAA TAAAGAAGGT GATATAATAG CTCTTGTTAC TTCATATACA
                                                                                           720
20
                                                                                           780
        GGATTTATGA GTGGAACCCT TGCGATCAGA AAAGCAGATG AAGGCAAATG GCAACTTGTT
                                                                                           840
        GGCGGAGATA TACAGAATGC GATCGTTCAA AATATATGCA TGATGGACGA CAACAAGATT
                                                                                           900
        GCTTGTGAAG TCTTCGGGAC TCCTAACGGA GTAGATGGTC GGACAAGGGT TTGTGTTTCT
                                                                                           960
        GACGCATCTG TCTTTGATTT TGAGTGGTAT GAAGATGAAA TATACGGAGG CCTGATATTT
                                                                                          1020
        GACACTTTCT TCTATAGCCC TTGGGACAAA CTTCTTTATG CGAAATTTGG TGGGATTATG
CTCAGGAGTA AAGAGTCTTT TATAACCTCT TTCATTTCTC CGACAGTTGT ACAAGGAGTG
25
                                                                                         1080
                                                                                         1140
        GATGTCTATA CTTTGGCCGG GAAGATAAGG ATCGAAAGTG AAACTCCGGT GTCTGAGGTG
                                                                                         1200
        TTGCTTTTCG ACCTGGCTGG CAGGATGGTA CTTCGGCAAA CCATTGATAA TAAAATCTAT
                                                                                         1260
        TCGGACATAG ATACTAACGG ACTAAAGCGA AGCGGTATTT ACGTAGTCTC GGTGCGGCTC
                                                                                         1320
30
        TCTTCCGGAC AGGTATTCAG TCATAAGGTG CAGGTA
                                                                                         1356
        (2) INFORMATION FOR SEQ ID NO:97
35
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 993 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                   (D) TOPOLOGY: circular
40
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
45
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
50
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...993
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97
55
       GGCCTGTACG GTATGTCTGT TGTACCTATA ATAATTTATC TTTGCGGTAT ATCAAATTAT
                                                                                           60
       GCAAGACTCA TGATAATCCG GTGTCTTATC CGTCGTCCGA GAACCGTCCT GTTCGGGTTG
                                                                                          120
       ATATTCGTGG TAGGTCTTTT CTCTGCGATG GCGCAAGAGA AAAAGGATAG TCTCTCTACG
                                                                                          180
       GTTCAGCCAG TGCCGAATAG CAGCATGGTG GAGCAGACCC CTCTTCTCTC CATTGATCAC
                                                                                          240
60
       CCCGTCCTGC CCGCTTCTTT TCAGAATACC CGTACACTGA AAAGGTTTAG AGACAAACAT
                                                                                          300
       CTTTCCGATG CTTTGCTCAA TGGATTGAAG CCTCATCGCT CATCTTTGCA ATTGAATGAG
                                                                                          360
       GAACTCAACT TCGCGGCAGA GCGTCGGGAT TTCGTTTCTC CCCTCTTGCA AACTCGCCAC GCTGCCGGTG TCCTTTCATG GCGACCGACC GATAGGATGC ATTTTTATAC ATCGGGCAAT
                                                                                          420
                                                                                          480
       ATCGGTCTTG GCCATGATTT ATTGACCGGT GTGCGCAAGG ACTTCGGATG GAATGCTGGT
                                                                                          540
65
       GCCGACTTCT TGCTGAGTCA AAATCTTACG GCACATGTCC AAGGCGGTTG GCAGCAGAAT
       TTCGGCTTTA TACCTATGAC GGCTGTCAAT GGCCAACTGC GTTGGCAAGC CACCGAGAGA
TTGAGTTTTA CCACCGGTAT CGATTATCGA CAGGTACAGT GGAATGCTTT CGATAATAGA
                                                                                          660
                                                                                          720
       ACGTTCTCGC TTAAAGGAAG TGCTCGATAC GAAGTGATGG ACAATGTCTT TGTCAATGGA
                                                                                          780
       TTTGGCAGCT ATCCTCTCTA CAGCAGTACG CGCTCAGGAC TCAATATGGC TGTTCCGATG
                                                                                          840
70
       CATGGATTCG GCCCTCAGTA CGGTGGATCG CTTGAGCTGA AAGTCTCCGA GCGATTCGGC
                                                                                          900
       TTTGCCGTCG GTATGGAGCG CGAATACAAT ATCTGGACTC GTCGGTGGGA AACGCATTAC
                                                                                          960
       TTTGCTTATC CTGTATTCTA TGGCGATAAG AAG
                                                                                          993
75
       (2) INFORMATION FOR SEQ ID NO:98
```

```
(i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 987 base pairs
                 (B) TYPE: nucleic acid
 5
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
10
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
15
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...987
20
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98
      GAGACGAACT CTTGGGTATC CAGCGATTGC AATTCGACGA CGATGAAAAC GAATAGACGA
                                                                                60
      TACGCATTTG TTTTGCCGCT TCTGCTACTC ACCGGATTGT TGGCATGGGG GCAGGATTCT
                                                                               120
25
      TCCCACGGTA GCAATACAGC GTTTGCAACT GATTCTTCGA GTAGAGAGTT GCCCACGGAG
                                                                               180
      CAGTCCGCCT ACCGCATTCA TTCTGCCTAT ATGGTCGGTG GTGGCGGAAG CATAACGCGC
                                                                               240
      GACACCTATT TGTCACCCCT TCGTTATGGA GGATGGACAC TGAATTTGTT GGGAGAGAAG
                                                                               300
      ACGTTCCCTC TCAAAGCCTC CGATTCCCGT TGGATGATCC GTACCGGGCA TGAGCTGGAT
                                                                               360
      TTTGCCCTGA TGGACAATCC GGCCAATAAT GCTCATTTCT ATTCCCTGCT GTATAACGGT
                                                                               420
30
      TCCGCTGCGG CTCTTTACCG CCTTGGCGCT AAGCATCTGC GAGCCGCGTG GATGGACAAT
                                                                               480
      CTGCGCTTGG CATTCGGCCC GGGCTTGGAA ATCGGGCTTG GAGGAATTTA TAGTACACGC
                                                                               540
      AACGGCAATA ATCCTGCGAC ATTGAAGCTC TACACCAATG CCATCGCCCA AGCCTCGATA
                                                                               600
      GGATACTACG TCCCCTCCGA AACTTTTCCC CTGTATTTTC GGTTGCTCTC CCAGATCAAT
                                                                               660
      CTCTTCGGTA TAGCCTATGG AAATGGTTTT GGTGAGAGCT ATTACGAGAA TTTTTTGCTC
                                                                               720
35
      AATAACGGCA TTGCAGGCTC CCTGCATTTC ACTTATCCGG GCAAGTTTAC TCGGTTCACG
                                                                               780
      ACACTCATAA CGGCGGATAT TCCCATTCGG AACTTCTGTA CGCTTCGTGT CGGTTATCGC
                                                                               840
      TATTCCCATT TGGGCTCTTC GCTTAACGCA TTGGATACTC GAATCCACAG TCATACGGCT
                                                                               900
      TTTATCGGTT TCGTCACGGA GTTTTACCGA TTCCGTGGGC GCAAAGCCAT GAATACCGGT
                                                                               960
      CGGAGAACCA GTCTTTACTA TCATGAT
                                                                               987
40
      (2) INFORMATION FOR SEQ ID NO:99
           (i) SEQUENCE CHARACTERISTICS:
45
                (A) LENGTH: 957 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
50
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) AUTI-SENSE: NO
55
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
60
                (A) NAME/KEY: misc_feature
                (B) LOCATION 1...957
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99
65
      TTTACGTCCG GTACGATATT CGTCCGTATA GCCATATCCG GCCGTGTAGT GTGCCGTGAG
                                                                                60
      GTTGAGGATA ACGGAAGGAG AGAAGCTGTG CGTCATGATG GCATGGTAGT GACGCTGCTC
                                                                               120
      GTAATTGTCG GTATTGTGGT AGTAGTGCGG TACTCCTTGC GCGTCCACGT ACATAAGACC GGCACTGTTG TATCGGCGGC CATATTTGGC TTCATCCTCC TTGGAAAGAC CGTTCCATGC
                                                                               180
                                                                               240
      GATACCOGTA ACTTCTTTC CTCCGAAAGT GATGAACCTG AGAGCCGTGT TGCTACCGAA
                                                                               300
70
      ATAGCCCACC TGTGCGAAAT AGGATTTCAG ATCCACGCTT CCTCTATCCA CGTAGCCGTC
                                                                               360
      CGAACCGATT TTGGACAGGC GGGCATCCAC TGCCCAATGG CGACCGATGC GTCCGCTACC
                                                                               420
      GAGTTTGACC GATCGGCGGA ATGTGCCGAA CGAACCTCCG CTCAAATCGA CACGGCCATA
                                                                               480
      AGGAGCCAGT CCCAAATTAT CCGTACGCAT ATTGACACTT GCCCCAAAAG CTCCGGCACC
                                                                               540
      ATTGGTGGAA GTACCCACAC CTCGCTGCAC CTGAAGGTCT TCGATGGAAG AGGCGAAGTC
                                                                               600
75
      GGGCATATTC ACCCAAAAGA CGGACTGAGA TTCGGAGTCG TTGAGGGGTA CTCCATTGGT
                                                                               660
```



```
AGTTATGTTG ATGCGATTGG CATCGGTGCC ACGCACGCGA AAGCCGGAAT ATCCGATACC
                                                                               720
       CGTACCGGCA TCGCTGGTGG CTACCACGGA GGGAGTCAGC ATCAGCAGAT AGGGGATGTC
                                                                               780
       ACGACCATAA TTGGACTTGG AAAGTTCGGC CTTGCGAACG TTGGTGTAAG CGACAGGGGT
                                                                              840
       TTTCGCCGTG GCGCGAGTAG CTACGACCTG TACGGTCTGG AGCTGCACAT TGCTAAGACT
                                                                              900
  5
       ATCTATCTCG CTGTTGGAGA CGGGTGCTTG TGCCGTCAGG CAGAAAGGCA GGACGGC
                                                                              957
       (2) INFORMATION FOR SEQ ID NO:100
10
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1842 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
15
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
20
           (iv) AHTI-SEHSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
25
           (in) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1842
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100
30
       CCATATAATG TCCAATCTAT TAGCAACAAG ACGATTAAAA AACAAATGGA AAACTTAAAG
                                                                               60
       AACATTCAGC CCAGAGAGGA TTTCAACTGG GAAGAGTTTG AGGCCGGTGG CGTCCATGCT
                                                                              120
       GCCGTGAGTC GTCAGGAGCA GGAAGCTGCT TATGACAAAA CGCTCAATAC CATCAAGGAA
                                                                              180
       AAGGAAGTGG TAATGGGTAG GGTAACTGCT ATCAACAAGC GTGAAGTGGT TATCAATGTA
                                                                              240
35
       GGGTACAAAT CGGAAGGTGT GGTACCTGCA ACAGAATTCC GCTACAATCC CGAACTCAAA
                                                                              300
      GTGGGAGACG AAGTGGAAGT TTATATCGAG AATCAGGAAG ATAAGAAGGG CCAGCTCGTC
TTGTCTCACC GCAAGGGTCG TGCCGCTCGC TCTTGGGAGC GCGTGAACGA GGCTCTCGAA
                                                                              360
                                                                              420
      AAAGACGAAA TCGTAAAGGG CTATGTGAAG TGTCGTACCA AGGGTGGTAT GATCGTCGAT
                                                                              480
      GTATTCGGTA TCGAGGCTTT CCTCCCGGGA TCACAGATCG ACGTGCGCCC CATTCGCGAC
                                                                              540
40
      TACGATGCAT TCGTTGAGAA GACGATGGAG TTCAAGATTG TGAAAATCAA TCAAGAATAT
                                                                              600
      AAGAATGTAG TTGTTTCCCA CAAGGTGCTC ATCGAAGCAG AGCTCGAACA ACAGAAGAAA
                                                                              660
      GAAATCATCG GCAAGCTCGA AAAAGGGCAG GTACTCGAAG GTATCGTCAA GAATATTACT
                                                                              720
      TCTTACGGAG TATTTATCGA CCTCGGTGGA GTGGATGGTC TTATCCATAT CACTGACCTT
                                                                              780
      TCATGGGGTC GTGTGGCTCA TCCGGAAGAA ATCGTACAGC TGGATCAGAA GATCAATGTC
                                                                              840
45
      GTTATCCTCG ACTTTGATGA AGATCGCAAG CGTATCGCTC TCGGACTCAA ACAGCTGATG
                                                                              900
      CCTCATCCTT GGGATGCTCT CGACAGCGAG CTTAAGGTAG GCGATAAGGT GAAGGGTAAA
                                                                              960
      GTTGTGGTGA TGGCAGATTA CGGTGCTTTC GTTGAGATTG CACAGGGCGT TGAGGGTCTT
                                                                             1020
      ATCCACGTAA GCGAAATGTC ATGGACACAG CACTTGCGTT CTGCTCAGGA CTTCCTGCAT
                                                                             1080
      GTAGGCGACG AAGTGGAAGC CGTGATCCTG ACGCTCGACC GCGAAGAACG CAAAATGTCG
50
      CTCGGTCTGA AGCAACTCAA GCCGGATCCT TGGGCTGATA TCGAAACTCG TTTCCCTGTA
                                                                             1200
      1260
                                                                             1320
      CACCCCAGCG AGTTTACGGA AGTAGGTGCT GATATCGAAG TTCAGGTAAT CGAGATCGAC
                                                                             1380
      AAGGAAAACC GTCGTCTCAG CTTGGGTCAC AAACAGTTGG AAGAGAATCC TTGGGATGTA
                                                                             1440
55
      TTCGAGACGG TATTCACTGT AGGATCTATC CACGAAGGAA CGGTAATCGA AGTGATGGAC
                                                                            1500
      AAGGGTGCTG TCGTTTCTCT GCCTTACGGT GTGGAAGGTT TTGCCACTCC GAAGCACATG
                                                                            1560
      GTGAAGGAAG ATGGCTCACA GGCTGTACTC GAAGAGAAGT TACCTTTCAA GGTTATTGAG
                                                                            1620
      TTCAATAAGG ATGCCAAGCG AATCATTGTA TCTCATAGCC GTGTATTCGA AGATGAGCAG
                                                                            1680
      AAAATGGCTC AGCGTGAAGC CAATGCAGAG CGTAAGGCTG AAGCCAAAGC GGCTCAGAAA
                                                                            1740
60
      GAAGCTGCTG CCGAAGCTGC CAATCCTGCA CAGGCTGTAG AGAAAGCCAC TCTCGGAGAC
                                                                            1800
      CTCGGCGAGC TGGCCGCTTT GAAAGAAAAG CTTTCAGAAA AC
                                                                             1842
      (2) INFORMATION FOR SEO ID NO:101
65
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 729 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
70
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
75
```

WO 99/29870

```
(iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
  5
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...729
10
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101
       TCGACAATAA TGAAAAAAGC TATTCTTTCC GGAGCGGCCT TGCTCCTCGG CCTATGTGCC AACGCACAAA ACGTGCAGTT GCACTACGAT TTCGGTCATT CCATCTACGA CGAACTAGAT
                                                                                      60
                                                                                     120
       GGACGTCCCA AACTGACTAC CACAGTGGAA AACTTCACAC CCGACAAATG GGGAAGCACC
                                                                                     180
15
       TTCTTCTTCA TCGACATGGA TTACACGGGC AAGGGTATCC AGTCGGCCTA TTGGGAGATT
                                                                                     240
       TCGCGCGAAC TGAAGTTTTG GCAAGCTCCC GTTTCCATTC ATTTGGAGTA CAACGGAGGC
                                                                                     300
       CTCTCCACAA GCTTTACTTT CGGACACGAT GCTCTAATCG GTGCCACCTA CACCTACAAC
                                                                                     360
       AACCCCTCCT TTACACGTGG ATTTACGATC ACGCCCATGT ACAAGCATCT GGGTGCGCAC
                                                                                     420
       GACTTCCACA CCTATCAGAT CACCGGCACT TGGTACATGC ACTTTCTGGA CGGTCTGCTT
                                                                                     480
20
       ACCTTCAACG GCTTCCTCGA TCTTTGGGGT TTCCCCCCAAG AGAACCCAAT CGGGGGCCCT
       GTGCTCAAAG AAGGGGATAA GTTCGTATTC CTGTCCGAAC CGCAGTTCTG GATCAACCTC
                                                                                     600
       AATCGCATCA AAGGCATCGA CAAGGATTTC AATCTCAGCA TAGGGACAGA GATGGAAATC
                                                                                     660
       AGCAGGAACT TCGCTCGCAT GGACAAATTC TCCTGCATCC CTACTCTTGC GGTCAAATGG
                                                                                     720
       ACTTTCAAC
                                                                                     729
25
       (2) INFORMATION FOR SEQ ID NO:102
             (i) SEQUENCE CHARACTERISTICS:
30
                  (A) LENGTH: 705 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
35
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
40
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
45
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...705
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102
50
       ATGAAAACAA TTAGTAAGAA CCATGCGGCA CGGATCTGTG CCGCCATTGC TTTGTTTGCA
       GTGTGTAACG GCCGGATAGC TGCTCAGGAT TTTCTCTATG AAATAGGAGG AGGTTTTGGT
                                                                                     120
       GCTGCTCAGT ATTTTGGCGA TGCAAACAGA GGCTTGTTCG GTTCATCCGG AGTAGGTTTG
                                                                                     180
       GAGTTGGTCG GACGTTACAA TTATAATTTT CGCTGGGCTT TCAGTACCAT GTTGGATTGG
                                                                                     240
       CGTACATTGA GAGGCGATAC CGATAAGTCC GGGAATGTCT TCCCCGATTT TGCTCAAGCG
                                                                                     300
      GATTITAAGG TCGGCTTGAC TCAGCTCCAC GTTAGAAGCG AATTTAACTT TCTCCCTTAT
AGCGATGGCT ATAAGTATCT TGGTACAGCT CGGCTGTCTC CTTATGTAGC GGCCGGGTTG
55
                                                                                     360
                                                                                     420
      TCTTTGGGTT TTGCTTCGGG TGCTAAAGGT TCGGCTTTTG CTCCCGGGAT TACTGCGGGA ATGGGAGTGA AGTATAAGCT TAAACCGCGG ATCAATGTCG GTATCGAGTA TTCTTTCACG
                                                                                     480
                                                                                     540
       GGGTTACTTA CCGATGCTTT GGATGCGCTG ACGGATAAAA GTGTTTGGCT CGAGGATCCA
                                                                                     600
60
       TATAAGATCA ATGACTCCTG GGTCAAAAAC AAGGATGCTA CAGGGGCCTT AGTGCTTAGG
                                                                                     660
      ATTACGTATG ACTTCGGCCT GCGTAAGACT TTTTGTAATA AACAA
       (2) INFORMATION FOR SEQ ID NO:103
65
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1308 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
70
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DMA (genomic)
          (iii) HYPOTHETICAL: NO
75
```



| | (iv) AHTI-SENSE: NO | |
|------------|--|--------------|
| | (vi) ORIGINAL SOURCE: | |
| 5 | (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11308</pre> | |
| 10 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103 | |
| | AATATAATGT ATAAAGACTA TAAGGGTTTG TATGCGTCGC TTCGGTGGTA TGCCCTGATC | 60 |
| | ATTGGGTTGC TATTTGCAGC AGACGGTATA CAGGCTCAGA ACAACAACTT TACCGAGTCG | 120 |
| 15 | CCTTACACTC GCTTCGGCCT TGGCCGTCTC GGAGAACGGA CGACTATTAG TGGGCATTCC | 180 |
| 10 | ATGGGAGGAC TCGGCGTCGG TCTGCGTCAG GGCACATACG TCAATGCCGT CAATCCTGCT TCATACTCGG CTGTGGATTC GATGACGTTT ATCTTCGATT TCGGTGCATC TACCGGAATT | 240 300 |
| | ACGTGGTATG CCGAGAACGG GAAAAAGGAC AATAGGAAAA TGGGAAACAT TGAGTATTTC | 360 |
| | GCCATGCTTT TTCCTATTTC CAAATCCATT GCTATGAGTG CGGGAGTGCT TCCTTACTCC | 420 |
| 20 | GCATCCGGGT ACCAGTTCGG ATCCGTTGAT CAAGTGGAAG GAGGCAGCGT CCAGTACACC CGTAAATACT TGGGGACAGG CAATCTGAAC GATCTCTATG TCGGTATAGG TGCAACCCCG | 480 |
| | TTCAAAAACT TCTCAATAGG AGCCAATGCT TCATCCCTTT TTGGGCGATT CACACAGC | 540 600 |
| | AGGCAGGTAA TCTTCTCCAC GGAGGCTCCT TACAATCCCG TACATCTCTC GACGCTGTAC | 660 |
| | TTGAAGGCTG CCAAGTTCGA CTTCGGTATG CAGTATCACC TTCTTCTCAA ATCAGATCGT TCGCTCGTTA TCGGTGCCGT CTATTCTCCG CGGGTGAAGA TGCATAGCGA GCTGACTCAG | 720 |
| 25 | ATAAAGAATC AGGTTCAGAA CGGTGTAGTA GTGGAGAGA TGCATAGCGA GCTGACTCAG | 780 840 |
| | GGAATGGACT ATTATACCCT GCCTCATACA TTGGGGATAG GTTTTTCTTA TGAAAAGAAA | 900 |
| | GATAAACTTC TCTTAGGAGC AGACGTCCAA TATAGTAAAT GGAAAGGCGA GAAATTTTAT | 960 |
| | AAATCCGATT GCAAATTCCA GGACAGAATA CGGGTATCTC TCGGCGGAGA GATCATACCG GATATAAATG CCGTTGGGAT GTGGCCTAAA GTTCGCTATC GCTTCGGTTT ACATGGTGAA | 1020 1080 |
| 30 | AATTCTTACC TGAAAGTGCC GACTAAAGGC GGTGTATATC AAGGATACCA TATCGTAGGT | 1140 |
| | GCTGTATTCG GTATAGGAAT CCCGCTCAAT GACAGACGTT CGTTCGTAAA TGTCTCTTT | 1200 |
| | GAATATGACC GATTGATCCC GAAGGAGGGT ATGATCAAAG AAAATGCTCT GAAATTGACC TTCGGCCTCA CGTTCAACGA GTCATGGTTT AAAAAGCTGA AACTGAAC | 1260 1308 |
| 25 | 11 11 11 11 11 11 11 11 11 11 11 11 11 | 1308 |
| 35 | (2) INFORMATION FOR SEQ ID NO:104 | |
| | 12) INTOMINITION FOR SEQ ID NO.104 | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| 40 | (A) LENGTH: 2835 base pairs (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: double | |
| | (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 4 5 | - | |
| | (iii) HYPOTHETICAL: NO | |
| | (iv) AHTI-SEHSE: NO | |
| 50 | (vi) ORIGINAL SOURCE: | |
| | (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| | (ix) FEATURE: | |
| 55 | (A) NAME/KEY: misc_feature | |
| 55 | (B) LOCATION 12835 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104 | |
| | AATACCATTT CGGAGAATGG GAACGATAGC ACCTATTTTT CTTTTCAAAC TTTTCTTATG | 60 |
| 60 | CGATCGATTT ATCAATTACT GTTGTCAATA CTCCTTGCTT CTCTTGGTTT CGTCGGGCTG | 120 |
| | GAAGCCCAAC AAGCCGGAGT AGCAGGTAGA GTATTGGACG AAGAAGGCAA CCCCATGATT | 180 |
| | CAAGCCAACG TACAGCTTGT ACAGAGTACC GGCCAAGTAG CCGTTGCCGC AGGTGCCACT AATGAAAAAG GGTTGTTCAG CCTGAAAACG TCACAGGAGG GTGACTACAT TCTGCGCGTT | 240 300 |
| CE | TCATATGTAG GTTACACTAC CCACGACGAA AAAATATCTC TTAGAAACGG GCAAACCATT | 360 |
| 65 | ACGCTCAAAG ATATATCCAT GAACGAAGAT GCCCGTCTTC TACAGAGTGT GACGGTGCAG | 420 |
| | GCTAAAGCGG CAGAGGTCGT GGTACGCAAC GATACGCTCG AATTCAATGC CGGATCCTAT ACCGTAGCAC AGGGAGCTTC TATCGAGGAA CTGATCAAGA AGCTACCCGG AGCAGAGATC | 480 |
| | GGATCCGATG GGAAGATCAC CATCAACGGC AAGGACATTA GCAAGATCCT TGTCCATGGC | 540 600 |
| 70 | AAAGAGTTTT TCTCCAAAGA TCCACAGGTG GCAATAAAGA ATCTTCCGGC CGATATGGTC | 660 |
| 70 | AATAAAGTAC AGGTACTGAA CAAACTGAGC GAGCTGTCGC GGATGAGCGG TTTCGATGAT GGAGAAGAGG AGACCGTAAT CAACCTGACG GTGAAGCCCG AAAAAAAGAA AGGCCTCTTC | 720 |
| | GGAACGCTTC AGGCCGGCTA CGGTACCGAC CAACGCTATA TGGCCGGAGG GAACGTCAAT | 780 840 |
| | CGGTTCGATG GAAATAAGCA ATGGACATTG ATCGGTAGTG CGAACAATAC GAACAATATC | 900 |
| <i>7</i> 5 | GGCTTTAGCG AGATGGACAG CGAGATGGGA TCCATGACCT TCTTCTCTCC CCAAGGCGGT | 960 |
| | GGTCGACGCG GCTTCGGCAA TAGTGGAGGT GTTACGTCTT CGTCGATGCT GGGCGGCAAC | 1020 |

WO 99/29870 PCT/AU98/01023

```
TTCAGTGTCG AATTCTCCTC TGCCCTTAAT ACAGGAGGCG ATGCACGCTA CGGATACAAC
                                                                                      1080
        GACAAGGCCA TAGAGACGAC CAAACGCGTG GAAAATATCC TCGCCGAAGG GAATACTTAT
                                                                                      1140
        ATGGACGAAA ATATATTGGA ACGCTCTTTC TCTCACAATG GTCAGGCGCG ATTTAGGATG
                                                                                      1200
        CAATGGAAAC CGTCCGAACG TACCGAAGTG GTATTCGAGC CGGATCTTTC GATATCCAAG
                                                                                      1260
  5
        ATCGATGGGT TCTTTAACGA CACATACGAG ACGAAAGATG CCACCGGAAT CTCTATCAAC
                                                                                      1320
        AAAGGTTCTA TCCACCAAAC TACACAAGGA AACAACTTCA GACTGAACGG AGAATTGGAT
                                                                                      1380
        ATCAGTCACA AGCTCAACGA CGAAGGCCGT ACGATCAGTG CCTCCGTCAG TGGCGGTCTG
                                                                                      1440
        ACCGACGAAG ACGGAGATGG CATATATCAG GCTGTGCTCC AAAGCGTGGA GACGAATCAA
        AAGCAATTCA ACGACAACTC CAACCTGCAA TATCGGCTTC GCCTCTCGTA TGTGGAACCG
                                                                                      1560
 10
        TTGGGTAAAA ACTACTTCGC ACAAGCGATT CTGAACAGAC GTTTCTCCCG TCGCAATTCG
                                                                                      1620
        GATCGTGAGG TGTACCGACT GGGCGATGAC GGGCAATACT CCATATTAGA CAGTCAGTAC
                                                                                      1680
        GGACTCTCCT ACAGTAACGA GTTCACCCAG TATCGCATCG GACTCAACCT CAAGAAGATT
                                                                                      1740
        GCCAAAACGT GGGACTACAC CGTAGGATTC AATGTGGATC CCAACAGAAC TGTCAGCTAT
                                                                                      1800
        CGGAGCGTAG CCGGAGTAGA GCAGGACAAA CTGGCTTTCA ATCGTGTCAA TCTCTCCCG
                                                                                      1860
        ATGCTCCGAA TCAACTACAA ACCGAGCAGG ACTACCAACC TCCGAGTGGA CTACCGAGGA
CGCACGACAC AACCATCCAT CAATCAGATC GCTCCCGTTC AGGACATCAC GAATCCGCTA
 15
                                                                                      1920
                                                                                      1980
        TTCGTGACGG AAGGCAATCC CGGTCTGAAG CCGAGCTATT CCAACAATGT GATGGCCATG
                                                                                      2040
        TTCTCGGACT TCGATGCCAA AAGTCAGCGA GCTTTCAACA TTGTTTTCTT CGGCAACTAT
                                                                                      2100
       ACATTCGACG ACATCGTCCC CAATACGCAC TACGATCCGT CTACAGGGAT CCGTACCACT CGTTACGAAA ACGCCTCCGG TACGTGGCAA GCGAATCTTC ATGGGACACT ATCGCTTCCA
                                                                                      2160
 20
                                                                                      2220
       CTCAAGAACA GGGCATTTTC TTTCAGGATG TCCTTGTTCA ACAGGTTGGC CGAAGGACAA
AGCTTCATCA ATGACGATAA GAACAAAGCT CTCTCTTTCC GAACGAGGGA ACGCCTGACG
        CTGACCTATC GCAACAATTG GATCGATACG AGTATCGGTG GCAATATCGG ATTCTATATG
                                                                                      2400
       GCGAATAATA GTCTGAGCGG ACAGAAAGAT TCTCGCACAT ACGATTTTGG CGGCAATTAT
                                                                                      2460
25
       CAAGTTGCCC TAACGCTTCC CTATGGATTC CGTATCGACA GCGATGTTGA ATACAATACG
                                                                                      2520
       AACTCCGGTT ACAGCGGAGG ATTCAGTCTG GACGAATGGC TTTGGAATGC TTCGCTTTCA
                                                                                      2580
       TACAGCTTCC TCCGTGACAA GGCCGGTACA CTGCGTGTCA ATGGCTATGA CATCCTCGGT
                                                                                      2640
       CAGCGGTCAA GTATCAGCCG TTCTGCTTCG GCCATCAATA TAGAAGAGAG CATGTCCAAT
                                                                                      2700
       ACGATCGGAC GCTACGTGAT GGTGGACTTT ATCTACCGAT TCAACGCCTT CAGTGGTGGT
                                                                                      2760
30
       GGATCTCGCA GCGATCATCA GCGTGGCAAT ATGAATCGTC CGGGCCCACC TTTCGGCGGT
                                                                                      2820
       GGCAGACGAC CGTCC
                                                                                      2835
        (2) INFORMATION FOR SEQ ID NO:105
35
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1236 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
40
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
45
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
50
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...1236
55
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105
       GGAGAGTATC CTGCAAACAG TAACGACAAA AAAGAGATGG TTATGAAGCT GATTAAAAGA
       AGTTTGCTCC TGCTTGGAGC GGTACTGCTG ATTACGCTTC CTGCGTACTC GCAGAATGAT
                                                                                      120
       GACATCTTCG AAGATGACAT CTATACATCG CGAAAAGAAA TACGTAAACA AAACCAAGTT
                                                                                      180
60
       AAAGACTGGC AAAACCAAGA GGACGGATAC GGCGACGATA CGGAATATAC AGTGGCTTCC
                                                                                      240
       GATCGGGACA TTGACGCCTA CAATCGTAGA GATGGCCAGT CCTACGATGG GAAAAAGTTG
                                                                                      300
       TCCAAAGACA AGAAAAGAGA CTCCACTCGT TCTTCTGTTC CCGGTCGCTA TAGTCGCCGC
TTGGCTCGAT TCTATAAGCC GAATACGATC GTCATTTCAG GTGCCGACAA TGTATATGTA
                                                                                      360
                                                                                      420
      ACTGATGATG GTGAGTATTT CGTCTATGGA GACGAATACT ATGATGACGC GTCGTCTGTA AACATTTACA TCAACAGTCC TTGGTGCGAT CCGTTCCCTT ATACGTCATG GTATCCATCT
                                                                                      480
65
                                                                                      540
      TTCTCCGGCT GGTACAACTA TACGTGGAAC TATCCATGGT TCTACTACGG TAGCCATATC GGATGGGGCG GTTATTACCC CGGATATAAT TGGTATTGGA GCTACTACTA TGATCCTTTC
                                                                                      600
                                                                                      660
      TACAATCCCT ATGGAATCGG TATGGGTTGG GGATATCCTT ATGGCTGGGG CAGCTATTAC
                                                                                      720
      GGTTGGGGTG GCTATCCGGG AGTGATACAT CACTACCACC ACTACCCCAA GAAGACCTAT
                                                                                      780
70
      TCCAATGGTC AGCATTCCGG AGCTTACTAT TCTTATGGCC GACCGAATCG TATCAAAGGT
                                                                                      840
      GGAACGTCCG GTGCCAAACT TGGGACAGGA CGCTACGATA GAATTCAAAA TTCGTCTTCG
                                                                                      900
      CAAAAAAATA AGTTCGGATT GCAGTCGAAC AAACCCAATA ATAATCTGCA AAATGTCAAG
                                                                                      960
      TCGGGACGTA CCGGCCGAGC CAATAGAGAC CGAAATATAG AAACGGTAAC TCCAAACAAC
                                                                                     1020
      GGGCAAAAGC AGAATCGTCC CGTATTCCAG CAGAATCAGT CCGGCAATGA CCGACCGACC
                                                                                     1080
75
      GGACGGAATA TCCGCAGCGA GAGACAGGGG GAAAATAACG ATAGGACATT TTCGACTCCT
                                                                                     1140
```



| | TCTCGTAGCA ATAGTAACGG TGGCTTCTCC ACGCCTTCTC GCTCTTCTC CGGCTCTATG AGCGGAGGTG GCGGACGTAG TGGCCGGGGA CGCAAT | 1200 1236 |
|------------|--|--------------------------------------|
| 5 | (2) INFORMATION FOR SEQ ID NO:106 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1803 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 15 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 25 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11803</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:106 | |
| 30 | AGTAATAGCA GCTCCCACAA GTGGTTAATT TATTATCATA TAGAAAAGAC TAAAAGTATT ATGATCCGCA AGTTGATCTT ACTGCTTGCT CTGATGCCGG TAGCCTCTGT GGCTTTTGCG GTACCAACGG ACAGCACGGA ATCGAAAGAC AATCGTATCC TTACAAGCAT GCAATCCTCC TCTTTGAATA GGGATGATGC TCCGGATAAA TGGCAACCTA TGCATGCCAA TTTCAGTATT | 60 120 180 240 |
| 35 | CAGAGCGATA TGCTGCTTTC TACTGCCCAA AAGTCCAAGA ACACCTGGTT CGGCAACTCC TATATCATGG GTATAATCAA GAACAATTAT CTGGAGTTTG GTGCCCGTTT CGAGGATCTC TATAAGCCCC TGCCCGGACA TGAACCCGAG ATGGGGCGTG GCGTTCCTCA CATGTATGTG AAGGGAAGCT ATCATTGGGC GGAGCTGACT TCTACGATCA GTTCGGTAGC GGTATGGTAT TCCGCACCTA TGAAGAGCGC AACCTCGGTA TAGACAACGC GGTTCCGCGC GGACGTATAG TACTCACTCC TTTTGATGGA GTGCGTGTCA AGGGTATTGC AGGACAGCAG | 300 360 420 480 540 |
| 40 | CGTAACTACT TCGACCGCAC GGGCAAGGTA TTCAATTCCG GCCGAGGCTA CCTACTGGGT TCTGATCTGG AGCTGAATGT AGAGCGTTGG AGCAGTGCCA TGGCGACAA TGACTATCAT TTGGCTATCG GGGGATCGTT CGTTTCCAAA CACGAAGCAG ACGAAGATAT ATTTGTGGT GTAGGCGAAG ATCGCAAGCG ACTCAACCTG CCGCTCAATG TCCCGATTAT GGGCCTGCGC | 600 660 720 780 840 |
| 4 5 | ACCAACTTTC AAAAAGGAGG TCTCGCCCTC TACGCAGAGT ATGGATACAA ATACAACGAT CCCTCGGCAG ACAATGACTA TATCTACCAC GACGGACAGG CTGCACTCCT CTCTGCCTCA TACTCCAAAA AAGGGATGAG TATCCTGTTG CAGGCCAAAC GTTGTGAGAA CTTTGCTTTC CGCAGCAAGC GAAGTGCCCA GCTCACACCG CTTATGATCA ACTATATGCC GGCTTTTACC CAAGCTCACA CTTATACGCT GGCGGCCATC TACCCCTATG CTACTCAGCC TCAGGGAGAA | 900 960 1020 1080 1140 |
| 50 | TGGGCTTTCC AAGGTGAACT GCGTTACAAC TTTGCTCGCC GGACAGCTCT CGGTGGACGC TACGGTACCG GCTTGCGTAT CAACGTTTCG CATGTGCGTG GTCTGGACAA AAAGATGCTC AAAGAGAATC CCGACGAACT GATCGGAACG GATGGCTACA CCGTTTCTTT CTTCGGCATG GGCGACCTCT ATTATTCGGA TATAGATGTG GAGATTACTA AAAAGGTAAG CCCAGGATTC AACTTTACGC TCACCTACTT GAATCAGATC TACAATAACA AGGTACTGCA CGGTGCAGCC | 1200 1260 1320 1380 1440 |
| 55 | GGAGAGAGC CTGAGAAGAT CTATGCCAAT ATCTTCGTCT ATGATGGTAA GTATAAGCTG AGTAATAAGG TAGCCCTCCG TACCGAACTG CAATATTTGC ACACGAAGCA GGATCAGGGT GACTGGATCT ACGGCATGGC CGAGCTCTCT ATCCTGCCTT CTCTGATGCT TTCCCTCTCG GAGCAGTATA ATATCGGAGA GACCAAGAAA CATTATGTCA TGGGGTCTGT CACCTATACT CACGGAGGACA ATCGAGTAGC TTTCTCTGCA GGCAAAACCC GTGCAGGGAT GAACTGCTCG | 1500 1560 1620 1680 1740 |
| | GGAGGTGTAT GTCGTGTGGT CCCTGAGACT CAGGGATTCT ACCTTTCTTA TAGCACCAAT CTG | 1800 1803 |
| 60 | (2) INFORMATION FOR SEQ ID NO:107 | |
| 65 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 756 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 70 | (ii) MOLECULE TYPE: DNA (genomic) | |
| 70 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SEMSE: NO | |
| 7 5 | (vi) ORIGINAL SOURCE: | |

WO 99/29870 PCT/AU98/01023

```
(A) ORGANISH: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
  5
                  (B) LOCATION 1...756
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107
       AGGGGGTCGA GTTCGGGCAT TTCAGCCCGT GGACGCGACA TGCGTTCTTT ATTTTTGAGC
 10
       GCGTTGCGCA GCTCCTCTC CCATGGTTCA GAGCGACGCA GTCGGATAAG TTCTTCTGTA
                                                                                   120
       GTCATGTCAA TAAGGCAGAA AATAAGGTTA TTCCATCTCT CGGTATGCGC CCAAACGCAT
                                                                                   180
       GATCATCTCA TOGAAATOCA CTTGGTGTGC ATCGAATTCG GGGCCATCGA CACAGACGAA
                                                                                   240
       TTTCGTCTGT CCTCCCACGC TTATACGACA AGCCCCACAC ATACCGGTGC CATCCACCAT
       AATTGTATTG AGAGAAGCTA TGGTCGGTAT CTCGTAACGT TTGGTCAGGA GAGAAACGAA
CTTCATCATC ACAGCCGGCC CGATCGTAAC GCAGAGGTCT ACCGTTTCCC GTTTGATAAC
                                                                                   360
 15
                                                                                   420
       GCTTTCCACT CCATCCGTTA CGAGGCCTTT CGTCCCATAA GACCCATCGT CTGTCATGAT
                                                                                   480
       GATCACTTCA TCGCTATTGG CTCGCATTTG TTCTTCAAGG ATAACCAGAT CTTTAGTTCT
                                                                                   540
       GGCAGCCAAT ACGACAATTA CACGGTTGCC TGCTTTGTGG AAAGCCTCCA CGATCGGGAG
                                                                                   600
       CAAAGGAGCC ACACCCACAC CGCCTCCGGC ACAAACCACT GTGCCGACCT TTTCGATATG
                                                                                   660
 20
       CGTACTCTGT CCCAGCGGAC CTACCACATC CGTGATATAG TCGCCGACTT CGAGTTCGGC
                                                                                  720
       CAATTTCTTG GAAGATTTGC CCACGGCCTG AACCAC
                                                                                  756
       (2) INFORMATION FOR SEQ ID NO:108
25
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2370 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
30
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
35
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
40
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2370
45
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:108
       CTCTTGTTTT CTTCTCCTTT ACCCCGAATG GATCGTCCTA AGCCTTCATA TATTGTTCGA
       ATAGCAGCCA TTCTCTGCTT GTTTGTCGGC AGGCCTTTGT TTGCGCAGAG CTATGTGGAC
                                                                                  120
       TACGTCGATC CGCTGATCGG GACGCTAAGT TCTTTTGAGC TGAGTGCGGG CAATACCTAT
                                                                                  180
50
       CCGGTGATCG GTTTACCGTG GGGAATGAAT AGCTGGACAC CGATGACCGG TGTACCCGGT
                                                                                  240
       GACGGCTGGC AATATACCTA CTCGGCACAC AAGATTCGCG GATTCAAACA GACCCACCAA
                                                                                  300
       CCCAGTCCTT GGATCAACGA CTACGGCCAA TTCTCCCTTC TTCCCCTTAC GGCACCGCAG
                                                                                  360
       AAGCCATCAT CGAACGACTC CATAGCTCTG ACTAAATGGT GCAAGCAACT CTTTTCGGAC
                                                                                  420
      GAACAGACCT CGTGGTTCTC GCACAAAGCG GAGACGGCGA CGCCATACTA TTATAGTGTC
                                                                                  480
55
       TATTTGGCCG ATTACGACAC ACGCGTGGAG ATGGCTCCGA CCGAGCGTGC AGCTATCTTT
                                                                                  540
      CGCATACGTT ATTCCGGCAA TACCGAAAGT GGCTCCGGTC GATGGCTTCG TCTTGATGCC
                                                                                  600
       TTTACCGGTG GTTCGGAGAT TAGCATCGTG GATCCTCACA CCGTAGTGGG CATATCTCGC
                                                                                  660
      AAGAATAGCG GAGGTGTGCC GGCTAACTTC GCCTGTTATT TCATCCTGCA GTCCGATACT
                                                                                  720
       CCTATGGCCG ATGTCCTGCT TGAGACAGAT ACCGGCAAGT CAGACGAAGG CACAAGGGCA
                                                                                  780
60
      TGGGCAGCCT GTCGCTTCGA TTCGCAAGAA GTTACCGTCC GGGTGGCATC TTCTTTTATC
                                                                                  840
      AGTGTCGAGC AGGCCGAAAG AAATCTTGCG GAAGTCAAAG GGCAGAGTTT CGACCGGATC
                                                                                  900
      AGACTTGCCG GTCGCGAAGC TTGGAATAAG GTGCTCGGAC GCATACATGT GGAAGGAGGA
ACGAAGGATG AGCGCACTAC ATTCTATTCC GCACTCTATC GCTGTCTGCT TTTTCCCCGT
                                                                                  960
                                                                                 1020
      CGCTTCTATG AGGAGGATGC TTCCGGCAAT TTTGTGCATT ACAGCCCCTA CAATGGAGAG
65
      GTACTTCCCG GTTATCTCTA TACCGATACC GGATTTTGGG ACACTTTTCG AGCCCTTTTC
                                                                                 1140
      CCCCTGCTCA ATCTGCTGTA TCCCGATGAA AACATTAAAA TTCAGGAAGG TCTGCTGAAT GTATATCGCG AGAGTGGCTT TTTCCCCGAA TGGGCCAGTC CGGGCCATCG GGATTGTATG
                                                                                 1200
                                                                                 1260
      ATAGGCAACA ACTCTGCTTC TGTTCTGGCG GATGCCTACC TCAAGGGTGT TCGGGTAGAA
                                                                                 1320
      GATACCCGTA CACTGATGAA CGGACTCTTG CATGCTACGA AAGCCGTCCA TCCGAAAATC
                                                                                 1380
70
      TCCTCCACGG GTCGCAAAGG TTGGGAGTGG TACAACTCCT TAGGTTATGT TCCGGCTGAT
                                                                                 1440
      GCAGGCATCG ACGAAAGTGC TGCCCGTACG CTCGAATATG CTTATAACGA TTGGTGCATC
                                                                                 1500
      CTCCGACTGG GGCGCACATT GGGTTGGGAT AGAGCTGCAT TGGACACGTT GGCTCATCGT
                                                                                 1560
      TCGATGAACT ATCGTCATCT GTTCGATCCG GAAACCAAAC TCATGCGCGG TAGAAATCAG
                                                                                1620
      GATGGTAGTT TCCGGACACC TTTTTCCCCT TTCAAATGGG GAGATGTATT CACGGAGGGC
                                                                                 1680
75
      AATGCCTGGC ACTACACTTG GTCGGTCTTT CATGATGTGC AGGGGCTTAT CGACCTGATG
                                                                                1740
```



| 5 | GGAGGAGATC GCCCGTTCGT GTCTATGCTC GATTCGGTAT TCAATACTCC TCCTATGTTC GATGAGAGCT ATTACGGATT TGTCATCCAC GAAATCAGAG AGATGCAAAT AGCGGATATG GGCAATTATG CTCATGGCAA TCAACCCATA CAGCATATGA TATATCTGTA TAATCATGCC GGTCATCCAT GGAAAGCTCA GGAGAGACTA CGCGAAGTGA TGGGGCGGCT CTATCGTCCT ACTCCGGATG GGTATTGCGG CGATGAAGAC AACGGACAGA CTTCGGCTTG GTACGTTTTC TCTGCTTTACA GCTTCTATCC TGTTACACCC GCTACGGATC AGTATGTGCT CGGTTCGCCG ATTTTTTCCA AGGTAATACT CTCTTTTCCC GACGGACACA AACGGTGTT GCATGCTCCG | 1800 1860 1920 1980 2040 2100 2160 |
|------------|---|--|
| 10 | GCCAACAGTG CCGATACGCC TTACATCCGC TCGATCAGCG TAGAAGGAAA AGAATGGAGC TGCAATTACC TGACTCACGA ACAGCTTCGC TCTTCTGCAT CCATTCAATG GATGATGGAC ACGAAACCCA ATTATAATCG TGGTATGAAG GAAAGTGACA GACCTTATTC CTTCTCCACG GAGCAACAGC GTCGCGCTAA TCACAGTAAT | 2220 2280 2340 2370 |
| 15 | (2) INFORMATION FOR SEQ ID NO:109 | |
| | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 858 base pairs | |
| 20 | (B) TYPE: nucleic acid (C) STRANDEDNESS: double | |
| 20 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 25 | (iv) AHTI-SEHSE: NO | |
| | (vi) ORIGINAL SOURCE: | |
| 30 | (A) ORGANISH: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: | |
| | (A) NAME/KEY: misc_feature (B) LOCATION 1858 | |
| 35 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109 | |
| 40 | ATTTGTGGCA GTAAAATGAA TCTATCGGGA TTACAATCTT TCACTATGAT GAAATCCATG CGCAGCGTGC TGTTGCTACT CTTTCCATTG TCTTTGATCA CTGCTTTGGG CTGTAGCAAT AACAAAGCTG CCGAATCGAA GTCTGTCTT TTCGATTCGG CCTATCTCGA ACGCTACATC CCTCTGCGGG CAGACATAGA TACGCCATCG CTGCATGTGA TGATCAGCTA CGTCTATCT CGCCGGAGAT ATATGCTCAC AGAGATTTTC AACGGTTTGC TCTTCGGCGA CAGCCTGATG CCCTTATCT CGCCGGAAAA TGCCATGGAA GGCTATGCAC AGAAGACTAT | 60 120 180 240 300 360 |
| 45 | CGCTCTAACA ATGCCGAAGC CAATCTGCAA GGGCTTCCTT CTGACCTTTT GGACTATATC TACAAGCAGG AAAATACCAT CGCTTATTGC GATACGGGAT TGATCTCCAC GCGCATCAAT ACGAAGGCGG TGACCATACG GAGAATACAG TCCGGTTTGC CAACATCCTT CGCACCACCG GCAAGGTGCT CGAAGAGCGA GATATATTCA AGGATCGACTA TGCGGAAAGG CTGTCCGCAC TCATCATAGG ACAATTGGTG CACGATTTCCA GCAAGACCAC ACCTGCCGAA TTGGATGCAA TAGGTTTCTT CAACGCAGAA GAAATACAGC CCAATGGCAA TTTTATGATC | 420 480 540 600 660 720 |
| 50 | GATGACAAAG GTCTCACATA CTGTTTCAAT GAGTATCAGA TAGCTGCTTA TGCCAGAGGT GCTGTCTATG TCCGTCTCGG ATATGACGTA TTGGCTCCTT TGCTAAGGGA TGATTCCCCA CTAAAGCGTT ACTTGCCG | 780 840 858 |
| 55 | (2) INFORMATION FOR SEQ ID NO:110 | |
| 00 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1134 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double | |
| 60 | (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 65 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SEMSE: NO (vi) ORIGINAL SOURCE: | |
| 70 | (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| | (ix) FEATURE: (A) NAM4E/KEY: misc_feature (B) LOCATION 11134 | |
| <i>7</i> 5 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110 | |

PCT/AU98/01023 WO 99/29870

| 5 | GGGATAATAT CTGTTCTTC ACATGTGGTT GACCGTCCAC AATGGGGTGC TTCTCCCGAA GCTGCTGGCA CGCATAGTGT GTATTCGATT CTACATCCT CCGCCGGTAT TATCCGGATC AGGTCTATGG GTATTATCTC AGCCTGCCGT ATCGCAATTC TTGCCGGCAA GCACCCCGGT CGCGGTCTG CGAGGGCAAA TGTAGGTATT CTTTCGTACA ATCCAGAAAA CACGCCCGAG AAAAAAAAAA | 60 120 180 240 300 360 |
|----|---|---------------------------------------|
| 10 | ATGCCCANT CGAAGTTAAA GATCAAGAGC ATTGAGGCAG CTACCACTTT CAGTTCGGCC ACGCCCGAA ATGGTTTTGG TGGCAATATC TTCGGCATGA ACATGAGCAT ACGGATGAGG GTACACCACA GCATTCTGCC CGAAGGGTTG GATTTTCGG TAGGAATACA TGAAAGAAGA GCACACTGGG AAGAGGCCG AAGTCCGAAG CTCATGTATA CGAATGTCCC AAGTATCATT GGTATTGTTG AAAAGGTAAT AGTCTTCGAA GACGCAGAAG ACTTTTTTGA CAAAAAAAGCT | 420 480 540 600 660 |
| 15 | CTCGGCCGCT TCCTCATCAG TTTGGGGATA TCCTTATACCA AGCATCTGGG AGCGTATTGG GGATGGACCA ATGACGCCA TATTCTTTC TCACCGATAC CCAAGAGCAA GGTCCACTAT GACACCTACA CAAGAGCTGG CAGTGACCTT GTACTTCAGT CCGAAGATGT TGCCACAGTG AGCAATGGCT TTTCACCGGG GATCGGACTC AAAAGTTCTA TTTGGTGGAA AATGCCCATC AAGAGCAAAT ATGATTTTCG CCTCGGTTTC AGCCTGGGCT ATGAGTATCT GAACCTGCTA | 720 780 840 900 960 |
| 20 | TATCCGTATC GTAATTTCAA GCTGGATGGA AATAAGCCGC TTTCAGCACT ATCTCCTCGC ATGAACCACA TCGGCCATGT GGGCTTCAAC TTTACCGTGG GTCTTTGGAC TAAT | 1020 1080 1134 |
| | (2) INFORMATION FOR SEQ ID NO:111 | |
| 25 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3807 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 30 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 35 | (iv) ANTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORTPHYROMONAS GINGIVALIS | |
| 40 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 13807</pre> | |
| 45 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111 | |
| | GTAAAATGTA TGGGCAAGTA TAAAAGAGCT AAGTACCGCT ATTGGCTTTT TCCTTTCTGT TCGGATTATT ATACCTTTGA GGGAGTTACT TTTTTATGCG CATCTGACGA TATGACAACC AAGAAAACCCC AAGCCATTTT AGACTTAGAG AAGGCCTATA ACATTGAAAT TCCTGATCTC | 60 120 180 |
| 50 | TCCTCACAAG AAGGGATAAG CTGGTCGGTA AATAGATATT TCAAGCAAGA TTCCTCCGGT GCAGTCGTTG AGCTTTGCTT GCGAGAATGT CAGATAGAAA GCATGACTTG GCTTATTGAT TTTCCTGCTC TAAAAAAGCT TGATCTATCG TATAACCAAA TCAGTAAGCT AGAGGGTCTA GAACGTCTTA CTTCGTTAAC AAAACTTCGT CTAAGAAGTA ACCAAATCCG TAAACTAGAG GGCCTGGATA GTCTCACCTC GCTAACAAAA CTTTCTCTCT CCGATAACCA AATCAGTAAG | 240 300 360 420 480 |
| 55 | CTAGAGGGTC TGGAACGTCT CACCTCGTTA GCGGAGCTTT ATCTTTTGGA TAACCAAATC AGTAAACTAG AGGGTCTGGA ACGTCTCACG TCCTTAGCAA CGCTTGAACT ATCGGTAAC | 540 600 |
| | CAAATCCGTA AGCTGGAGGG TCTGGAACGT CTCACGTCCT TAGCAACGCT TGAACTATCG GGTAACCAAA TCCGTAAGCT AGAGGGTCTG GAACGTCTCA CTTCGTTAAC AAAGCTTCGT | 660 720 |
| 00 | CTAAGAAGTA ACCAAATCAG TAAGCTAGAG GGTCTGGAAC GTCTCACGTC CTTAGCAACG CTTGAACTAT CGGGTAACCA AATCCGTAAG CTGGAGGGTC TGGAACGTCT CACGTCCTTA | 780 840 |
| 60 | GCAACGCTTG AACTGTCGGG TAACCAAATC AGTAAGCTAG AGGGTCTGGA ACGTCTCTCT TCGTTAACAA AGCTTCGTCT AAGAAGTAAC CAGATCAGTA AACTAGAGGG CCTGGAACGT CTCACCTCGC TAACAAAACT TTCTCTCTCC GATAACCAAA TCAGTAAGCT AGAGGGTCTG | 900 960 1020 |
| 65 | GAACGTCTCA CCTCGTTAGC GGAGCTTTAT CTTTTGGATA ACCAAATCCG TAAGCTGGAG GGCCTGGAAC GTCTCACCTC GTTAACAAAG CTTCGTCTAA GAAGTAACCA AATCAGTAAA CTAGAGGGCC TGGATAGTCT CACCTCGCTA ACAAAACTTT CTCTCTCCGA TAACCAAATC AGTAAACTAG AGGCCTGGA ACGTCTCACG TCCTTAGCGG AGCTTTATCT TTTGGATAAC CAAATCCGTA AGCTGGAGGG TCTTGATGGT CTTGCTTCCT TAACAAGGCT TAGTCTAAGG | 1080 1140 1200 1260 1320 |
| 70 | CGCAACCAAA TCAGTAAGCT GGAAGGACTA GACAGACTAA AGGTTTTGAG AAAACTTGAT GTTTCGGGCA ATGATATTCA ATCTATTGAT GATATTAAGC TATTGGCTCC GATTCTGGAG CAAACTTTAG AAAACTGAG AATCCATGAC AATCCATTTG TTGCATCATC AGGCTTGATA CTCTCTCCTT ATGATAATCA TTTGCCGGAG ATTAAAGCTC TTCTTTGAAAA AGAAAAAGAA AAACAGAAAA AGACTTCAGT TGAATATCAC CCATTTTTCCA AAGTAATGCT ATTGGGAAAT | 1380 1440 1500 1560 1620 |
| 75 | CATTCTTCGG GTAAAACAAC ATTTCTTAGT CAATACGATA CAAATTATAC GTATCAGAAA AATACACATG TGTTGTCGAT ACATCGAAGC AATAACCCTA ATGCGATCTT TTACGACTTT GGGGGACAGG ACTATTATCA TGGGATTTAC CAAGCCTTTT TTACCACCCA ATCGTTATAC | 1680 1740 1800 |



76 / 490

| | CTTCTCTTTT GGGATGCTAA GAAGGATCGA AACTTTGTGA GCGTAGATGA TAAAGAATAT | 1860 |
|------------|--|--------------|
| | CAGACTCTTA ATTTCAATCG CCCCTATTGG TTAGGACAGA TAGCCTATGC CTGCAATCGT | 1920 |
| | TGTATGTCCG TTGGAGGAAA TCCTGATGGC AAGGACACAC CACAGACCAC AGACGATACA ATTATCATTC AGACTCATGC CGATGAAACG GGCGCTAAGC AGCAAACCTT AGGCTGTGCA | 1980 |
| 5 | GCCGAGAATG GAGTATTGGA AGAAATCTAT GTATCCTTAG AGCCCAAGGC GAATAGTGCC | 2040 2100 |
| | GTACATGCGC TCAACTATCT GAATGAGCGG GTGCGAGAAG TTGTCGCAAG CAGGAGTAAA | 2160 |
| | TCAATTCAGA TCACAGAAAA AGATAAGGGA TTGTACGAAG CTCTTCCCAC AATCGCCGGT | 2220 |
| | GATAATAAAC ACATCCCTAT CTCTCTCGAA GCTCTTGCGG CTCAATTGAA TAAGGGAAGA | 2280 |
| 10 | GCTGAAAATG ATCTTTACAC CATAGAGTAT CTACAGACCG AATTGAACCA GCTTAGTCTG CGAGGGGAGG TGCTTTACTA TCGTGAGAAT GAGAAGCTGA ACAATTATGT CTGGTTAGAT | 2340 |
| | CCGGCAGCTT TTGTCCAAAT GATTCATGAG GAAATCCTCC AAAAAGACAA CATCAATAGA | 2400 2460 |
| | GGAACAGTTC CTAAAGACAT TTTTGAATGC AAACTGCATA ATCTAAGTTC CGGAAGTATA | 2520 |
| | TTTGAAGAAG ATGGCCAAAA TGGTAATATG ATCTTGCAGC TATTATTGGA AGAGCTGATC | 2580 |
| 15 | GTATATGAAG ATAAGGACTG CTATGTGATA CCGGGCTATC TCCCTTTGCA TTCCGATGAC | 2640 |
| 13 | GAAGCCTATA AATGGCTTAC TTTGGGATTC GAGAGGCCCA ATTTTGTCCT CAAATTCGAA | 2700 |
| | CGTTTTATCC CCTTTGGCCT GATCAACCAG ATTATAGCCT ACTATGGCCG GGAAGAAGGT GCTCTAAAGC GGTATTGGCG AGATCAGGTC ATCTTCACAG CAGGCCGTGA GATGGATAGG | 2760 |
| | CAAACGCTTG AGCAAGAAGA AGAGAAAGAG GGTTTGCCCA AGACGAATGC CGAGGATTAT | 2820 2880 |
| • | CAGATCTGGA TCAAGCTCGA CTTTACCGAC TTGGCCATAT CCGTATTCAT CAAAGAGCAG | 2940 |
| 20 | AGAAAGACAT CAGCTAAGGA TATGCAGCGG AAAGAGGCTA CTATCCTCAG TGATATGTTG | 3000 |
| | GATATGTATT GGAACAATAT CCCTCCGAGG GAGCAAATAG GAGATAAGGA TACGGAGCAA | 3060 |
| | ACGAGAAGCA CTATTCGTGA AACAAACAGA AAGAAGAGC CCATCCAGGA TCTCTACCTC TCCTGTGCCC AAGCGGATAA AGATTTGACG GAGTCTCATT ATATCCATTT GGGCACGCTG | 3120 |
| | GACGATGAAA GCAAGACTAC GGCGAGGATT GCAGCCTATC CGTTGAAGAA CGGCGTTATC | 3180 3240 |
| 25 | GATAAAGAGC GGGTGCGAGA AGTATCGACT CGTCCCTACA AACATCTTTC CGTCAATAAA | 3300 |
| | AATCTGGCTA CTGCAAAACA GATCTTTATT TCCTATTCCA AAGAGGATCA GACTGAACTG | 3360 |
| | GAGACCTGTC TGCAATTTTT CAAACCCTTG GAGAAGAATG GTCAGATCGA GATCTACTAT | 3420 |
| | GATAAGTTGA CTAAGTTTGA AACACCTATT CACCCTGAAA TAAGAAAGCG TATTGTCGAA | 3480 |
| 30 | GCCGACTGTA TAATCGCTTT GATCAGCCAA CGCTATCTGG CCACGGATTA CATCCTGGAT | 3540 |
| | CATGAGTTGC CTGTATTTCG GGAGTATAAC AAGACCATAG TGCCGATATT GATCAAGCCT TGTACATTCG AAGACGATGA GTTCCTTCGG GAGAAATATT TTGCTCAGAA AGCTCAAATA | 3600 |
| | ATCAATCTTG GAAAAGAGGG AAAAACCATT AAAGCTTATG ATAGTATTAC GGCATCAGCC | 3660 3720 |
| | CATCGTGATG AAAATTGGGT GGCAGTAGTC AGAGAGTTCA AAGAGAAGAT ATTAAGAATA | 3780 |
| 35 | ACAAAACAGG AGGTAAATAC AGATGAA | 3807 |
| 33 | | |
| | (2) INFORMATION FOR SEQ ID NO:112 | |
| | | |
| 40 | (i) SEQUENCE CHARACTERISTICS: | |
| 10 | (A) LENGTH: 693 base pairs (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: double | |
| | (D) TOPOLOGY: circular | |
| 45 | 1223 May Paul B. Burg. Burg. | |
| 43 | (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| | (iv) AHTI-SEHSE: NO | |
| 50 | . , | |
| | (vi) ORIGINAL SOURCE: | |
| | (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| | (ix) FEATURE: | |
| 55 | (A) NAME/KEY: misc feature | |
| | (B) LOCATION $16\overline{9}3$ | |
| | (vi) CECHENICE PROGRESSION OF THE PROGRESSION OF TH | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112 | |
| 60 | AAGTTTATGA TGAAAAAAGC ATTTGTTTTC GTACTACTGG TTTGCCTATT CTCCTCGTTC | 60 |
| | AGCAGTTCCG CCCAAACAAC GACGAACAGT AGCCGGAGTT ATTTTACAGG ACGAATCGAG | 120 |
| | AAGGTGAGTT TGAACTTAGG GGTCCCCCCC GTAAGCACAG AGGTTTGGGG AATGACCCAT | 180 |
| | GATGCGAACG GTCTCCCTTT CGAAATACCT ATCTCTTTCA GTCGTTTCAA CAGCCAGGGA | 240 |
| 65 | GATATAGCTA CCACTTATTA CATAGCGAAT AGCGAGGCAA CTTTGAATGA ATGGTGCGAC | 300 |
| | TATGCACACC CGGGCGGCAT CGTGAGGGTA GAAGGTCGTT TTTGGAAAAT GACTTACAAC ATACCAACCT ACAATGCAGT CTGCACCCGG ATTACATTCG AAAATCAAGA AATAGAAGGA | 360 |
| | ACGATCGTCT TGATACCCAA GCCCAAAGTC TCGCTGCCTC ATGTGTCGGA ATCGGTGCCT | 420 |
| | TGCATCCGAA CCGAAGCCGG GAGGGAATTT ATCCTTTGCG AAGAAGACGA CACCTTTGTG | 480 540 |
| 70 | TCTCACGATG GTAACGAAGT AACGATAGGC GGTAAACCTT TCTTGCTCAA TACCAACGTA | 600 |
| 70 | AAGATTGTGG GGGACGTATC TCAAAAGTAT GCCGTGGGGG TAGGAGAAAT TCGATTCCTG | 660 |
| | CAGATTTGTG CCCAAACAGT ATCACAACAA AAA | 693 |
| | | |
| | (2) INFORMATION FOR SEQ ID NO:113 | |
| <i>7</i> 5 | | |

SUBSTITUTE SHEET (Rule 26) (RO/AU)

PCT/AU98/01023 WO 99/29870

```
(i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1689 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
 5
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
10
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
15
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1689
20
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113
      TCGCTAATCA ACAATTCAAG AGACTGGAGG GCTCCTGTCC GATTGAGTCT CAAAAAAAAG
                                                                                 60
      ACAAAAACTA TGAAGACAAA AGTTTTACGC AAATTCGTGG TGGCGGCTTT CGCCGTCGCA
                                                                                120
      ACCCTCTGTC CTCTCGCCCA AGCGCAGACG ATGGGGAGGAG ATGATGTCAA GGTGGTCCAG
                                                                                180
25
      TACAATCAGG AAAAACTGGT ACAAACGAGG ATGAGTGTGG CGGACAACGG ATGGATCTAT
                                                                                240
      GTAATGACCC ACAGTGGATA CGACACCGGC AATAGCAATG TGAAGATCTT CCGCTCCAAA
                                                                                300
      GACCAAGGTG CCACATACCA AAAGTTGAGG GATTGGGATC CATCGGATGA TTATCAGTTT
                                                                                360
      CAAGACTTCG ATATCGTGGT AACGGGTAAG AATGAATCCG ACATCAAGAT TTGGTCGGTA
                                                                                420
      GAGCTCATGA ATAAGCCCGG AGGATATAAG AGTAGAGTTG CGGTCTTCAG TCGCGATGCC
                                                                                480
30
      AACGCGCAGA ATGCGAAACT CGTGTATAAG GAAGACTTCT CCAATGTGCA GTTGTACGAT
                                                                                540
      GTGGATATAG CCTCCAACTA TCGTTCGCCT TCTTCTCTTA ACAATGGTGG CAACCCTTTT
                                                                                600
      GCTTTGGCTT TCGCTTACAC CGGCTTCAAC AATACGCACA AAATAAGTTT TGTGGACTAT
                                                                                660
      GTGTTCTCTC TGAATGGAGG GCAAAATTTC AATAAAAACT TACTCTTCAG TCAAGATGGA
                                                                                720
      GAGAAGAAA TTGACAAGGT GGATCTCTCA TTGGGTAGCA CCTCTGAATC CATGGGTCAC
                                                                                780
35
      AATGCCTGGC CGCTAATGGG TGTGGTATTC GAAATGAATA AACAAGGGGG AAAAAGCGAT
                                                                                840
      ATCGGTTTCT TGTCGAACTT TGTCGACAAT GATCCCGAAT TTCAGTGGTC AGGCCCTATA AAAGTGAGTG AAAGCGACAT GTCGTTCAGC CCCAAAATCC AAATGTTGCT GGACGAGGAT
                                                                                900
                                                                                960
      AACAATACGA TCAATGGGGA GAGTTGCCAC AACTTCATGA TTACGTACAG CGATTATGAT
                                                                               1020
      TCTGAATATT CGGATTGGGA CATTCGGTAT GTATATCCCA AGAAATCGTT CAAGTATGAA
                                                                               1080
40
      AAAGGAAAAA CTCCGACTAT GGATGATCTG GTGGAAGCTT TCCTTACAGC TTCGTACCAG
                                                                               1140
      AGTGAGACCA ACTCGGGGCT GGGGTATGAC AAGAACGCCA ATCACTACCT GATTACATAT
                                                                               1200
      GCCAAAAAAG AAGAGAACGG TACGAACACG CTGAAATACC GCTGGGCCAA TTATGACAAG
                                                                               1260
      ATTCATAACA AAGATTTGTG GAGCGACACA TTTACGTATA CATCATCTGC CAATGCTCTC
                                                                               1320
      TACACACCTC AAGTAGACAT CAATCCGACC AAGGGTCTCG TGTGCTGGTC ATGGGTGGAA
                                                                              1380
45
      TATCTGCCGG GCAAACGGAT CGTTTGGTCT GATACGCAGT GGACCCATGC CAACGGTGTA
                                                                              1440
      GAAGACATCG TAATGCAAGA AGGCAGCATG AAGCTCTACC CGAATCCGGC TCAAGAATAT
                                                                              1500
      GCTGTGATTA GCCTGCCGAC GGCAGCAAAC TGCAAGGCTG TTGTTTACGA TATGCAGGGC
                                                                              1560
      AGAGTAGTCG CTGAGGCTTC TTTCTCCGGC AACGAATACA GGCTGAACGT GCAGCACTTG
                                                                              1620
      GCTAAGGGTA CGTACATACT CAAGGTCGTA TCCGATACGG AGCGTTTCGT AGAGAAGCTC
                                                                              1680
50
      ATCGTGGAA
                                                                              1689
      (2) INFORMATION FOR SEQ ID NO:114
55
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2358 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
60
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
65
          (iv) AUTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
70
          (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                (B) LOCATION 1...2358
          (XI) SEQUENCE DESCRIPTION: SEQ ID NO:114
75
```



| | ATCGTTTACC TTTGTCACTG TATGAACCAC AGACGATCAA AAACCATGCT GACGATCCGA | 60 |
|---|---|---|
| | AACTTCCTCC TCTTTTGTTG TCTGTCGCTG ATAGCGTTTG CTGCCGATGC ACAAAGCTCT | 120 |
| | GTCTCTTCGG GTAGACGACT GACAGAATAT GTCAATCCCT TTATCGGAAC GGCCAATTAC | 180 |
| - | GGTACCACGA ATCCGGGAGC AGTATTGCCC AATGGGTTGA TGAGCGTTAC CCCTTTCAAT | 240 |
| 5 | GTCAGCGGAT CGACAGAGAA TCGCTTCGAC AAAGATTCGC GTTGGTGGAG TGCGCCTTAT | 300 |
| | TCGGCCGACA ATAGTTACTG CATCGGTTTC AGCCATGTGA ATCTGAGTGG AGTAGGCTGT CCCGAACTGA GTGGAATACT GCTGATGGCC ACTTCCGGCA CATTCGATCC TGATTACTGC | 360 |
| | TGCTATGGCT CTTCGCTCAG TCGAGAATAT GCGCGCCCGG GAGAATACAA GGCTGTATTG | 420 480 |
| | GACAAATACG GTATAGATGC AGCCGTGACC GTAACCGAGC GGACTGCTTT GACCGAATTT | 540 |
| 10 | GCTTTTCCCG AAGGAGAAGG CCATATCCTG CTGAACCTGG GACAGGCCCT AAGCAATGAA | 600 |
| | TCGGGAGCCT CTGTTCGATT CTTAAACGAC TCCACAGTCG TCGGCAGCAG GCTGATGGGG | 660 |
| | ACGITCTGCT ACAATCCGCA AGCAGTTTTT CGTCAGTATT TCGTACTTCA GGTGAGTCGG | 720 |
| | CGACCGATCT CTGCCGGCTA TTGGAAGAAG CAGCCTCCTA TGACAGTGGA AGCCCAATGG GATTCGACTG CAGGGAAATA TAAGCAGTAC GACGGCTACA AGCGTGAGAT GAGCGGTGAT | 780 |
| 15 | GACATCGGTG TCCGATTCTC GTTCAACTGC GATCAGGGGG AAAAGATCTA TGTACGATCG | 840 900 |
| | GCCGTTTCAT TCGTCAGCGA AGCCAATGCG CTCTATAATC TGGAAGCGGA GCAAGAAGAG | 960 |
| | GTGTTCAAAA GTGTCGGAGG GAATCCGGCC AAGGCTTTCT CCGCTATACG CTCTCGCGCT | 1020 |
| | ATAGAGCGTT GGGAGGAAGC CCTCGGTACG GTGGAAGTGG AAGGAGGCAC ACCGGATGAA | 1080 |
| 20 | AAGACGATAT TCTATACCGC ACTCTATCAC CTGCTGATAC ATCCGAATAT CCTACAAGAT | 1140 |
| 20 | GCCAATGGAG AATATCCTAT GATGGGCAGT GGCAAAACGG GTAATACGGC TCACGACCGC TACACCGTGT TCTCTCTTTG GGACACGTAC CGCAATGTAC ACCCGCTGCT CTGCCTCCTC | 1200 |
| | TATCCGGAGA AGCAGTTGGA TATGGTACGG ACACTGATCG ACATGTACCG AGAGAGCGGG | 1260 |
| | TGGCTGCCGA GATGGGAGCT GTACGGACAG GAGACCCTGA CGATGGAGGG CGACCCCTCG | 1320 1380 |
| | CTTATCGTCA TCAATGACAC TTGGCAAAGG GGCCTTCGTG CTTTCGATAC GGCAACGGCC | 1440 |
| 25 | TATGAAGCCA TGAAAAAAA TGCTTCTTCG GCAGGAGCGA CCCATCCGAT CCGTCCTGAC | 1500 |
| | AACGACGACT ATCTCACCCT CGGCITCGTA CCGCTTCGCG AACAGTACGA CAATTCCGTA | 1560 |
| | TCGCATGCGC TGGAATACTA TCTGGCCGAC TGGAATCTGT CCCGGTTTGC CCACGCACTT | 1620 |
| | GGGCATAAAG AAGACGCAGC TCTATTCGGA AAACGCTCGT TGGGCTACAG ACACTATTAT AATAAGGAGT ATGGTATGCT GTGTCCATTG CTGCCGGATG GATCATTCCT CACTCCTTTC | 1680 |
| 30 | GATCCCAAAC AGGGTGAAAA CTTCGAGCCT AATCCCGGTT TCCACGAGGG CAGTGCTTAT | 1740 1800 |
| | AACTATGCCT TTTTCGTTCC CCACGATATA CAAGGGCTTG CCCGGCTGAT GGGAGGAGCA | 1860 |
| | AAGGTTTTTT CGGAAAGGTT GCAGAAAGTC TTCGATGAAG GATATTATGA TCCGACCAAC | 1920 |
| | GAGCOGGACA TOGCCTATCO TTACCTCTTC TCCTATTTCC CCAAGGAAGC ATGGCGAACG | 1980 |
| 35 | CAGAAATTGA CCCGGGAGTT GATAGACAAA CATTTTTGCA ATGCTCCTAA CGGCTTGCCC GGTAATGACG ATGCCGGTAC GATGAGTGCT TGGCTTGTCT ATTCCATGCT GGGATTCTAC | 2040 |
| 00 | COTGACTGTC CGGGCAGCCC CACCTATACA CTGACCTCGC CGGTATTCCC CCGAGTTAGG | 2100 |
| | ATTCGGCTCA ATCCGCAGTA TTATCCTCAG GGGGAGTTGA TCATTACGAC CAATACAGAG | 2160 2220 |
| | AATCAACCGA CAGATTCCAT TTACATCCAT ACGGTTTCTC TTGGCAATAA AACACTTCCG | 2280 |
| 40 | CATGGAACAA GGCATATCAG CCATGCCGAT TTGGTGCGCT GCGGTCACCT CCGTTACGAA | 2340 |
| 40 | CTAAGCAATC GTCCTCGA | 2358 |
| | | |
| | (2) INFORMATION FOR SEQ ID NO:115 | |
| | (2) Intolution for BEQ 15 NO.115 | |
| 45 | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 2442 base pairs | |
| | (B) TYPE: nucleic acid | |
| | | |
| | (C) STRANDEDNESS: double | |
| 50 | (C) STRANDEDNESS: doubl∈ (D) TOPOLOGY: circular | |
| 50 | (D) TOPOLOGY: circular | |
| 50 | | |
| 50 | (D) TOPOLOGY: circular | |
| | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO | |
| 50 55 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) | |
| | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO | |
| | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) AHTI-SENSE: NO (vi) ORIGINAL SOURCE: | |
| 55 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: | |
| 55 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) AHTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc_feature | |
| 55 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: | |
| 55 60 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: | |
| 55 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: | |
| 55 60 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: | 60 |
| 55 60 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: | 120 |
| 55 60 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12442 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115 CCCCGGTTTTC ATCCCGGAGG GACACACGCT CAGCTTTGCA GAAATGGGAG AAGAAACCAA AAATCAAATC AGTCATCGTG CTTTGGCGGT GGCACAGCTC CGCGATTTTT TATTATGTGC AAAATAAGAT TCAGCCTCTT GCCAGGCTCTT GTCGTCTGCT TATTGTTCAC CTCTTTTTCT | 120 180 |
| 55 60 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12442 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115 CCCCGTTTTC ATCCCGGAGG GACACACGCT CAGCTTTGCA GAAATGGGAG AAGAAACCAA AAATCAAATC AGTCATCGTG CTTTGGCGGT GGCACAGCTC CGCGATTTT TATTATGTGC AAAATTAAGAT TCAGCCTCTT GCAGGCTCTT GTCGTCTGCT TATTGTTCAC CTCTTTTTCT CTCCAAGCTC AAGAGGAAGG TATTTGGAAT ACCCTCTTGG CTATCCACAA GACGGAAAAA | 120 180 240 |
| 556065 | (II) MOLECULE TYPE: DNA (genomic) (III) MOLECULE TYPE: DNA (genomic) (III) HYPOTHETICAL: NO (IV) ANTI-SENSE: NO (VI) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (IX) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12442 (XI) SEQUENCE DESCRIPTION: SEQ ID NO:115 CCCCGTTTTC ATCCCGGAGG GACACACGCT CAGCTTTGCA GAAATGGGAG AAGAAACCAA AAATCAAATC AGTCATCGTG CTTTGGCGGT GGCACAGCTC CGCGATTTT TATTATGTGC AAAATAAGAT TCAGCCTCTT GCAGGCTCTT GTCGTCTGCT TATTGTCAC CTCTTTTTCT CTCCAAGCTC AAGAGGAAGG TATTTGGAAT ACCCTCTTGG CTATCCACAA GACGGAAAAA CGCCTAGAAA CGCCCAAGAA AGTCTTTTGCC GTAGCCAACG GAGTACTTA CTCGGTGGGC AAAGAAGCTC CCCATGAGGC AAAGATCTTC GACCGTATCA GCGGACTCAG CGATCATCC | 120 180 240 300 |
| 556065 | (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12442 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115 CCCCGTTTTC ATCCCGGAGG GACACACGCT CAGCTTTGCA GAAATGGGAG AAGAAACCAA AAATCAAATC | 120 180 240 300 360 |
| 556065 | (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12442 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115 CCCCGTTTTC ATCCGGAGG GACACAGGT CAGGTTTGCA GAAATGGGAG AAGAAACCAA AAATCAAATC AGTCATCGTG CTTTGGCGGT GGCACAGCTC CGCGATTTTT TATTATGTGC AAAATAAGAT TCAGCCTCTT GCAGGCTCTT GTCGTCTGTT TATTGTTCAC CTCTTTTTCT CTCCAAGCTC AAGAGGAAGG TATTTGGAAT ACCCTCTTGG CTATCCACAA GACGGAAAAA GCCGTAGAAA CGCCCAAGAA AGTCTTTGCC GTAGCCACG GAGTACTTA CTCGGTGGGC AAAGAAGCTC CCCATGAGGC AAAGAATCTTC GACCGTATCA GCGGACTCAG CGATACATCG GTAAGCACA TAGCCTACTC CGAGCAACTA AAATCCTTGG TCATATACTA TGCATCAGGC AATATCGACA TCTTGGACGA AGCAGGCCGT GTGACCACACG TACCTGCATT GCATCAGGC AATATCGACA TCTTGGACGA AGCAGGCCGT GTGACCACACG TACCTGCATT GAAAGACAAT | 120 180 240 300 |
| 55606570 | (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12442 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115 CCCCGTTTTC ATCCGGAGG GACACACGCT CAGCTTTGCA GAAATGGGAG AAGAAACCAA AATCAAATC AGTCATCGTG CTTTGGCGT GGCACAGCTC CGCGATTTT TATTATGTGC AAAATAAGAT TCAGCCTCTT GCAGGCTCTT GTCGTCTGT TATTGTTCATCTCCCAAGCTC AAGAGAAGAA AGTCTTTGCC GTAGCCACAG GACGAACAA GCCGTAGAAA AGCCTCTTG CTATCCACAA GACGGAAAAA GCCGTAGAAA AGCTCTTGC GTAGCCACA GACGGAACAA GACGGAACAA AATCATCA TCTGGTGGGC AAAGAAGCTC CCCATGAGGC AAAGAATCTC GACCGTATCA GCGGACTCAG CGATACTCA GACAGACAA TCTTGGACGA AGCAGCCGT GTAGCCACG GATACATCG GTAAGCAACA TCTTGGACGA AGCAGCCGT GTAACCACG TACCTGCATT GAAAGAACAT ATCGATCTA TAGGACAACAA GCCCTTAATTTC | 120 180 240 300 360 420 |
| 556065 | (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12442 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115 CCCCGTTTTC ATCCGGAGG GACACAGGT CAGGTTTGCA GAAATGGGAG AAGAAACCAA AAATCAAATC AGTCATCGTG CTTTGGCGGT GGCACAGCTC CGCGATTTTT TATTATGTGC AAAATAAGAT TCAGCCTCTT GCAGGCTCTT GTCGTCTGTT TATTGTTCAC CTCTTTTTCT CTCCAAGCTC AAGAGGAAGG TATTTGGAAT ACCCTCTTGG CTATCCACAA GACGGAAAAA GCCGTAGAAA CGCCCAAGAA AGTCTTTGCC GTAGCCACG GAGTACTTA CTCGGTGGGC AAAGAAGCTC CCCATGAGGC AAAGAATCTTC GACCGTATCA GCGGACTCAG CGATACATCG GTAAGCACA TAGCCTACTC CGAGCAACTA AAATCCTTGG TCATATACTA TGCATCAGGC AATATCGACA TCTTGGACGA AGCAGGCCGT GTGACCACACG TACCTGCATT GCATCAGGC AATATCGACA TCTTGGACGA AGCAGGCCGT GTGACCACACG TACCTGCATT GAAAGACAAT | 120 180 240 300 360 420 480 |

PCT/AU98/01023 WO 99/29870

```
GCCAAGGGAA CTAAGGTGAC CGATGTGGCT AAGTTGGACA ATGATCGCTT GCTGATGCTG
AAAGAAGGGC AGCTCTTCAT CGGAAAAGAG ACCGATAACC TGCAAGATCC GGCCGCATGG
                                                                                         660
                                                                                         720
        ACAGCCTTGT CTTTGAATTT GCCGATGGGC TCGGTCACCG GTCTGGGCAT TGTCGGGGAA
                                                                                         780
        GACATCTGTT TCCTGCTCGC CGATGGCCGT GTATATGTCG CTGCAAACCA ATCGTTTGAG
                                                                                         840
        CCGGAGCTAT TGCTCTCTC CTCCGCCGAT TCACGACTGT ATGTGACGGA TCGTGGTCTG
  5
        TTCATCTGTG CCGAGAATCG AATTTATTTC ATAGAAAAAG GTCGCAAAAC GACACAATTT
                                                                                         960
        CCTATAGCCG ACGTCCTTGG TGTCGGTGCC ATGAACGAAA GCAATACGGC ATACATAGCA
TTGGGAGAAG AAGGTTTGGC TTCACTTCTT CTCGCAGAGG GAAGTACGGC CGAAGCCATG
                                                                                       1020
                                                                                       1080
        CCTGTAGCAT TCGACGGACC GGGGGACAAT GATTTCTACG AGATGCGGTT TAGTCACGGA
                                                                                       1140
 10
        CGTCTGTATG CAGCCAGCGG ACTCTGGGGA ACAAACCTGA TGGGACATGC CGGTATGGTG
                                                                                       1200
        AAGCTATACG ACGGCAACCG ATGGACTAAC TTCGACAAGA AGACCGTACA GGAACAGTTG
                                                                                       1260
        GGCGGCGGAT TCAGTTTCAA TGATGCTATC GATATAGCTG TTTCCAACGG AGACCCCGAT
                                                                                       1320
        CACTITITG TCGGTACATG GGGAAACGGT CTGTTCGAAT TCAAGGATGG CAAAGCGATA
                                                                                       1380
       GCTCGCTATT CGGGAAACGA AACTGCTATC GCAGAATGTA ATCCCGGAGA TGCCCGTGTG
AAAGCGATTG CCTTTGACAA TAAGGGCAAC CTCTGGGGGA CGCTCGGTGC CGTAGGCAAG
                                                                                       1440
 15
                                                                                       1500
        AACATCTTCA TGTACGATCC GCAGAGTAGC ACATGGCATT CTTTCAGCTA TCCGGATGTA
        GCCAATCTGG CCTCCTTCGG CAATATGATT ATCCTACCCA ACGGAGACAA ATGGGTAAAT
                                                                                       1620
        ATCCTTCACC GTAGTGGCGG ATCCACGCGC AAAGGTGTCT TGATCTTCAA CGATCGGGGT
                                                                                       1680
       ACACCGGAAA CGACTTCGGA CGACAGCCAT CTTTACGTCG AGCAGTTTGT CAATCGCCTC
                                                                                       1740
20
        GGGGCAGCCA TAGGACATAA GACTATCTAT GCAATGGCCG TCGATCATAA CGGCTCTGTC
                                                                                       1800
        TGGATGGGAT CGGATATAGG CATTTTCGGC GTCTACAATG CAGCCGGAGT ATTGTCCTCG
                                                                                       1860
        ACTICTACCC CTATCGCTGT TCGGCCGGTC GGAGGAGAAG AACCCAATTT GTACTATGTG
                                                                                       1920
        CTGGACAAGG TGACGGTGAC AGACATCGTC GTGGACAAAC TCAATCACAA ATGGGTTGCC
                                                                                       1980
        ACCCAAGGGA CAGGACTCTA TCTCCTTTCG GAAGATTGCA GTAAGATCCT CGCGCAATTT
                                                                                       2040
25
       ACCGTAGAAA ACAGCCCTTT GCTTTCTAAC AACATACTAT CCCTGGCCTT AAATGACGAT
AACGGACTGC TGTACATCGG TACGGCGGAC GGACTGATGA CGTTCCAAAC GGGTACGGGG
                                                                                       2100
                                                                                       2160
       AGTGGATCAG CTTCCGAACT GGACGGCGTC TATGTATACC CCAATCCGCT AAGGCCGGAA
TATCCCGATG GCGTCACCAT TGCCGGACTG CAAGCCGGCT GTAGTGTCAA AATCACCGAT
                                                                                       2220
                                                                                       2280
       ACCACCGGCA GACTGCTATA CCAGACTGAG AGCGTAACCA CCGAAGTCAA ATGGAATGCT
                                                                                       2340
30
       CGAGGTGCCG ATGGCAATAG GGTAGCTTCG GGCGTATATG CCGTTGCAGT GTACGATCCG
                                                                                       2400
       GTATCGAAAA AGTCCAAACT AATTCGCTTC GCAGTGATTC GC
                                                                                       2442
        (2) INFORMATION FOR SEQ ID NO:116
35
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LEHGTH: 3486 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
40
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
45
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
50
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...3486
55
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116
       GCTATTTCTC AGATGAAACG AATACTTCCA ATAGTCGCAT TCCTTTCTCT CTTCCTTGCC
       CTTGCTTTGC CTGCGAAAGC GCAACGAGCT ATGGGGAAGA CGGCCGACCG TTCGCTAATG
                                                                                        120
       GCTTCGGGAC ATTGGGTCAA GATACGTGTC GATGCAAGTG GAGTGTATCG CCTTACGGAC
                                                                                        180
60
       GAACAGCTCC GTGCCAATGG CTTCTCCGAT CCGTCCAAGG TAGGTGTGTT CGGTTATGGT
                                                                                        240
       GGAGGGGTGC TTCCCGAAGA TCTGAGCCGG ATCACGACAG ACGATTTGCC TCCGGTACCG
                                                                                        300
       GTACTCCGTC AGGGCAATGC GCTGTATTTC TATGCCGTGG GCCCGGTGAC ATGGTTCTAC
                                                                                        360
       AATCCGGCCA AAACCACCAT GGAGCATACG GTGAATACAT ACAGTACGCA TGGCTACTAC
                                                                                        420
       TTCCTGTCGG ATGCTGCCGG AGCACCTTTG CAGATGTCCC AATATACGGG TGGAGGTGCG
TCGGCCGAGG CTTTGATCGA CTACTACGAT GAGCTGATGC TCCATGAACA GGAATTGTAT
                                                                                        480
65
                                                                                        540
       TCGCCCAAAG AATCGGGACG AGATCTGTAT GGCGAGTCTT TCAGTGCAGT CAATACGCGT
ACGGTCAAGT TCCCTTTGAG GGGCAACACC CGCTCGTCTG GCGAACTCGG TACCGTATTC
                                                                                        600
                                                                                        660
       TCATACATAS CCAAGGCCAS ATCGGCCGGT GGCGGCCGTG AGATGTCGCT CTCGGCGAAT
                                                                                        720
       GGCATTCTGA TCTTCAGCGA TCCTTTTTCC ATGACATCGA ATGAAGTCTC CAATTCCTAT
70
       TTGGCCGGCA AGAAGCGTCG TCTCTATCAC AGTACGCCGA TGAACAGCTT GGTCAATGAG
                                                                                        840
       TTGCGCTTGG ACGCGAACTA TAGCATGACA GGAGATGCGG TCAATCTGGA TTTCATAGAG
                                                                                        900
       GTGGCTACAC AGAACGACCT CCGGTACGAT GGCGCACCCA TGCATATCAG GCGGTTTTCC
                                                                                       960
       AATTTGCCCG TTTTGGGGGG CGAGTCCTGC CGGTTCGTTA TCAGTGAGGT GCCGGAGTCT
                                                                                      1020
       CTGGTGGTTT TGCAGGCCAA TTCTTCCCTG ACAGCATCGC TTGTTCCCGT TAAGACTGTC
                                                                                      1080
75
       GGGGATAAGA CCATTGAGTT CGTGGCTCCG CCGAAGGGTC AGGATCGTAG GACTATCAAT
```



| | ACGTTTTATG | CCGTGGACTT | GTCACAGGCT | TCTGCTCCGG | AGATCCTCGG | AGCGGTACCC | 1200 |
|-----|------------|--------------|---------------------------|--------------|------------|------------|--------------|
| | AATCAAAACC | TGCATGGAGA | GGAAATCCCT | GATCTGATCA | TTGTCTCTAC | TCAGGCGCTC | 1260 |
| | CTCCTTGAGG | CTGATCGACT | GGCCACCTAT | CGTAGAGAGA | AAAACGGGCT | GAAGGTTTTG | 1320 |
| _ | | | GTTCAACGAG | | | | 1380 |
| 5 | | | GTTCTACGAC | | | | 1440 |
| | | | TCTCTTCGGT | | | | 1500 |
| | | | TCTCCAACAA | | | | 1560 |
| | | | TTATGTGACG | | | | 1620 |
| 10 | | | TTGGCGCAAT | | | | 1680 |
| 10 | | | CATCGCAGTG | | | | 1740 |
| | | | TCGTGCCTGT | | | | 1800 |
| | | | CGATACCGTC | | | | 1860 |
| | | | TCCGCATGTC | | | | 1920 1980 |
| 15 | | | AACCCTTCAG GTCGGACGAG | | | | 2040 |
| 10 | | | CATTTGGATT | | | | 2100 |
| | | | GGAGGAGGTT | | | | 2160 |
| | | | CGTTTACAAT | | | | 2220 |
| | | | AGCTAAGGAT | | | | 2280 |
| 20 | | | GCTCAGTACT | | | | 2340 |
| | | | GTCCGTGCGT | | | | 2400 |
| | ACCGCAATCA | ACGGGCAGGA | TCCCGAAGGG | CAGTATGGAA | CTATTATGCT | CAAGTCTTTG | 2460 |
| | GAACGGGTAG | CTCTGAAGGG | TAAGGTAACC | GATGAAAAGG | GGACATTCGA | CGAGACATTC | 2520 |
| ~= | | | CGTCTTCGAT | | | | 2580 |
| 25 | | | TGTATATTAT | | | | 2640 |
| | | | CTTCGAAACT | | | | 2700 |
| | | | CAATCTTTAT | | | | 2760 |
| | | | CAGAGTCCAA | | | | 2820 |
| 30 | | | AAGCTGCTTC | | | | 2880 2940 |
| 30 | | | GTTTATGGCC TGATATTACG | | | | 3000 |
| | | | CACAAGTTCG | | | | 3060 |
| | | | CGAAGGAGAT | | | | 3120 |
| | | | TGACTTTTCA | | | | 3180 |
| 35 | | | CCCGAATCCG | | | | 3240 |
| | | | TTTGAACGTG | | | | 3300 |
| | | | CAAGACCTAT | | | | 3360 |
| | AAGTGGGATC | TGACCTCCAA | ATACGGAGTG | AAGATCGGAA | ACGGATTCTA | CCTCTATCGT | 3420 |
| | TGTGTGGTGA | ACTCTCCCGG | AGGACAGACG | GCCTCCATGG | CCAAGAAAAT | GATCGTGGTA | 3480 |
| 40 | GGACAA | | | | | | 3486 |
| | | | | | | | |
| | | | | _ | | | |
| | (2) INFORM | ATION FOR S | EQ ID NO:11 | 7 | | | |
| 45 | (i) a | COURNER CUN | N COMP D T COM T CO | 3 . | | | |
| 70 | | | RACTERISTIC: 2919 base | | | | |
| | | | ucleic acid | Jairs | | | |
| | | | DNESS: doub | ا ۵ | | | |
| | | (D) TOPOLOG | | | | | |
| 50 | | , . , | | | | | |
| | (ii) H | OLECULE TYPE | E: DNA (gene | omic) | | | |
| | | | - | | | | |
| | (iii) H | YPOTHETICAL | : NO | | | | |
| | | | | | | | |
| 55 | (iv) Al | NTI-SENSE: 1 | 10 | | | | |
| | | | | | | | |
| | | RIGINAL SOU | | | | | |
| | | (A) ORGANISI | 4: PORYPHYRO | OMONAS GING. | IVALIS | | |
| 60 | (iv) F | EATURE: | | | | | |
| 00 | | | i: misc fear | rure | | | |
| | | (B) LOCATION | | curc | | | |
| | | (2, 200 | | | | | |
| | (xi) S | EQUENCE DES | CRIPTION: SI | EO ID NO:11 | 7 | | |
| 65 | | | | | | | |
| | | | GCATATAGAA | | | | 60 |
| | | | | | | TGTAACTGAA | 120 |
| | | | TATTACTGCA | | | | 180 |
| 70 | | | GAACAGCGTA | | | | 240 |
| 70 | | | ACGGATGCGT | | | | 300 |
| | | | CGCTGCATTG | | | | 360 |
| | | | GGATAGCCGT | | | | 420 |
| | | | CTTGATAGAA | | | | 480 |
| 75 | | | CCTGAACGGC | | | | 540 |
| / J | CATGATGAGG | CCGAAAAGAA | TTTCCTGAAA | GCACTGCAAG | GTGAGAAAGA | ACTCGACAGT | 600 |
| | | | | | | | |

WO 99/29870 PCT/AU98/01023

```
CCTCTCGGGC AAGCTATCAA CTATGCGAAC CTCGGACGTA TCTATCGACA ACGCAAAGAA TACGACAAGG CTCGTACCTA CTTCCTCTTG TCTCTGGAGC AGAACAATAT GGCAGAGAAT
                                                                                              660
                                                                                              720
        CTGATGGGTA TCGGACTCTG TAGCATCAAT CTCGGAGAAG TAGACGAAGA AAAAGGGGAT
TATCAAAAGG CTTTGCAAGA GTATGCCACG GCATACAAAC TGATGGAACA GTTGTCCGAT
                                                                                              780
                                                                                              840
        CGATGGCACT GGCTGAATTC CTGTATCCCG ATGGCACGTA TCAATCTCAA ACAAGGTAAC GAAAGGCTCT ACCAGCATTT CATTTCTTTG GCCGAAGGGA CTGCGAAAGA AATTAATTCG
                                                                                              900
                                                                                              960
        ACTTCACATC TGATAGAAAT ATACAATCTT CAATACGAGA ATCTCGAGCG TAAAAAAGAA
                                                                                            1020
        TACAAACAAG CCCTCGAAGC ATTCTGTCTG AGCAAGACGT TGAGCGACAG CATGTCCATT
                                                                                            1080
        GCGCACAAGG TCAGCAGCAT ACAAGAAACG CGATTCAACT ACGAACGAAA CAAGTCCCAA
                                                                                            1140
10
        AAAGAGCTTG AAGAAATACA GCAAGTAAGC AAGGCAAAAC AAGAGAAATC GAAGTTTATC
CTCTTGAGCA CTCTTTTGC CCTTTTCATC TCGATTCTTT TGATTTCTGT TCTGACATAT
                                                                                            1200
                                                                                            1260
       GCATACCGTC AGGGCAAGAA GCATAACAAG CTGATCAAAG AGACGGATAA ACTTCGCTCC
GGCTTTTTCA CCGGTATTAC ACACGAATTT CGTACGCCTA TCACCGTCAT ACAAGGTTTG
                                                                                            1380
        AATGAGAAAA TGAGTTCAAG TCCTGATCTC CAAGCATCGG ACAGAACCGA GCTGCACAAG
                                                                                            1440
15
        ATAATAGACA GACAGAGTAG CCATATGCTG AATTTGGTGA ACCAGCTGTT GGATATTTGC
                                                                                            1500
       AAGATCAGAA GCGGAGTATC CACGCCCGAA TGGCGCAATG GCGACATCGT CTCCTTCGTA CAGATTCTCA TCGATTCGTT TGCACCATAC GCACAGGCTC AAGACATAAC CTTGGAGCTA
                                                                                            1560
                                                                                            1620
        CAACCCGAGA GCAAACCTAT TGTCGTGGAC TTCGTCCCCT CCTACTTGCA AAAAATCATA
                                                                                            1680
        TCCAATCTTT TGTCCAATGC CATCAAGTAT TCTTTAGCCG GAGGGAGAGT GGTCATATCT
                                                                                            1740
20
        CTGGCAAAAA CCAAGAATGA AAAAAATCTG ATCATACGCG TTGCAGACAA TGGCATAGGA
                                                                                            1800
        ATAGATAAAA CTGATCAGGC TCATATCTTC GACATCTTCT ATCGAGGACA GTCCGCTACC
                                                                                            1860
        GAAAAGCATG GATCAGGCGT CGGACTCTCG TTTACCAATA TACTGGTCGA AAACCTTCGA
                                                                                            1920
       GGTACGATCA AAGTGGAAAG CCAGCCGGGG AAAGGAAGTG CCTTCACCAT CAGTATTCCT
ACACAAAACC AGTCCTCTTC GGCAGAGATT CTTCCTTGGC TACCCTCCTC CGATGACATT
                                                                                            1980
                                                                                            2040
       GTCATGCCTG TCCACATCGC GCCCGATGAC TCACCGACAT CTCCGATGGT AGCAGCTCTG
AATCATCGCT TCGAGGACGA ACGTCCGACC ATACTGCTCG TCGAGGACAA TAAGGATATC
25
                                                                                            2160
       AACCTGCTCG TCAAACTACT CCTTTGCGAT CGCTACAATG TGCTATCCGC CGCAAACGGA
AAAGAGGGTA TAGCCCTCGC TACCGAGCAT ATTCCCGACA TTATCATTAC GGATATTATG
                                                                                            2220
                                                                                            2280
       ATGCCGATAA TGGATGGGAT AGAAATGACA ATCCGGATGA AGCAATCGCC TCTGCTCTGT
                                                                                            2340
30
       CACATTCCCA TTGTCGCTTT GACGGCCAAG AGTACCGAAC AGGACAGATT GGAAGGAATC
                                                                                            2400
       AAAAGCGGTG TAGTCTCTTA TCTATGCAAG CCATTCTCTC CGGAGGAGCT TTTGATGCGG
ATCGAGCAGC TTCTGAAAGA CCGTGAGTTG CTCAAGAAGT TCTATATGCA AAAACTCATG
                                                                                            2460
                                                                                            2520
       CTGGATCGGA AGCCGGAGGA GGAGCCTCAA CCGATAGATG ACAGCAGTAT GCAGTTTCTC
                                                                                            2580
       CTTGCTGCCA AAGATGCAGT GTCCGGTGGA ATCAAACAAA ATCCGGATTT TTCCGCTCAA
                                                                                            2640
35
       GACTTGGCCG AAAAAATGTG CATGAGTCCA TCCCAACTCA ACAGAAAGCT CACGAGTGTC
                                                                                            2700
       GTAGGTTGCT CCACCATCGG CTACATACAG CAGATCAAGA TAAAATTGGC CTGCAAGCTC
                                                                                            2760
       CTTGCCGATG AGAGCAAAAA CATCTCCGAC ATTAGCATTG AGGCAGGCTT TTCGGATCCG
                                                                                            2820
       GCTTACTTCT CTCGCACCTT CAAACGCTAC ATGAACTGCT CTCCCTCCCA ATATCGGCAA
                                                                                            2880
       AAACTCCTTG CCATGCCGGG GAGCGACAAG GAGACAGTT
                                                                                            2919
40
       (2) INFORMATION FOR SEQ ID NO:118
              (i) SEQUENCE CHARACTERISTICS:
45
                    (A) LENGTH: 1689 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
                    (D) TOPOLOGY: circular
50
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
55
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
60
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...1689
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:118
65
       CATCATAAAA CATATCAAAC AATGAAAAAG CTTTTACAGG CTAAAGCCTT GATTCTGGCA
       TTGGGACTCT TCCAACTGCC CGCAATCGCC CAAACGCAAA TGCAAGCAGA CCGAACAAAC
                                                                                             120
       GGTCAATTTG CAACAGAAGA GATGCAACGA GCATTCCAGG AAACGAATCC CCCTGCAGGT
       CCTGTGCGTG CTATCGCTGA GTACGAACGC TCTGCAGCCG TTTTGGTACG CTACCCGTTC
                                                                                             240
       GGTATCCCGA TGGAATTGAT CAAAGAGCTG GCCAAGAACG ACAAGGTGAT TACCATTGTG
                                                                                             300
70
       GCGAGTGAAA GCCAAAAAAA CACCGTTATA ACCCAGTACA CCCAAAGCGG TGTGAATCTC
                                                                                             360
       TCTAATTGCG ATTTCATCAT TGCGAAAACT GACTCTTACT GGACACGCGA CTATACCGGT
                                                                                             420
       TGGTTCGCAA TGTACGATAC GAACAAAGTA GGTCTCGTGG ACTTTATTTA TAACCGCCCT
                                                                                             480
       CGTCCTAACG ATGATGAATT CCCCAAATAC GAAGCACAAT ATCTGGGCAT CGAGATGTTC
                                                                                             540
       GGGATGAAGC TCAAGCAGAC CGGTGGCAAC TACATGACGG ACGGATATGG ATCCGCTGTG
                                                                                             600
75
       CAGTCACATA TCGCATATAC GGAGAACTCC TCTCTGTCTC AAGCTCAAGT AAATCAAAAG
                                                                                            660
```



```
ATGAAAGACT ATCTCGGCAT CACACATCAT GATGTGGTAC AAGATCCGAA CGGCGAATAT
                                                                                  720
       ATCAACCATG TGGACTGTTG GGGCAAGTAT TTGGCACCGA ACAAAATCCT CATCAGGAAA
GTGCCTGACA ATCACCCTCA GCACCAAGCC CTGGAAGATA TGGCAGCCTA CTTCGCAGCA
                                                                                  780
                                                                                  840
        CAGACCTGCG CATGGGGAAC GAAGTACGAG GTATATCGCG CTTTGGCCAC CAATGAACAA
                                                                                  900
  5
       CCGTACACGA ACTCTCTGAT TCTGAACAAC AGGGTATTTG TTCCTGTCAA TGGCCCCGCC
                                                                                  960
       TCCGTGGACA ACGATGCTCT GAACGTCTAT AAGACGGCAA TGCCCGGTTA CGAAATTATA
                                                                                 1020
       GGTGTCAAAG GGGCTTCAGG AACACCTTGG TTAGGAACAG ATGCCCTGCA TTGTCGTACT
                                                                                 1080
       CACGAGGTAG CGGATAAGGG CTATCTCTAT ATCAAGCACT ACCCGATACT GGGCGAACAG
                                                                                 1140
       GCAGGCCCTG ATTATAAGAT CGAAGCAGAT GTCGTCTCAT GCGCCAATGC TACTATCTCG
CCGGTACAAT GTTACTATCG TATCAATGGT TCCGGTAGCT TTAAGGCTGC TGATATGACG
                                                                                 1200
10
                                                                                1260
       ATGGAATCAA CAGGTCACTA TACTTATAGC TTTACAGGTC TTAACAAGAA TGATAAGGTA
                                                                                1320
       GAATACTATA TCTCTGCCGC TGACAATAGT GGTCGCAAAG AGACTTATCC CTTTATCGGC
                                                                                1380
       GAACCTGATC CTTTCAAGTT TACGTGTATG AACGAAACCA ATACATGTAC TGTGACCGGA
                                                                                1440
       GCTGCCAAAG CTCTTCGTGC ATGGTTCAAC GCCGGTCGTT CAGAACTGGC TGTTTCGGTA
                                                                                1500
15
       AGTTTGAATA TTGCCGGCAC ATATCGGATA AAGCTTTATA ACACCGCAGG AGAAGAAGTC
                                                                                1560
       GCTGCAATGA CCAAGGAATT AGTAGCAGGG ACGAGTGTCT TCAGTATGGA TGTGTATTCT
                                                                                1620
       CAGGCTCCGG GCACATATGT TCTGGTTGTT GAAGGAAATG GAATCCGTGA GACAATGAAA
                                                                                 1680
       ATTCTCAAA
                                                                                1689
20
       (2) INFORMATION FOR SEQ ID NO:119
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1311 base pairs
25
                  (B) TYPE: nucleic acid
                  (C) STRANDEDMESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
30
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
35
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) MAME/KEY: misc feature
40
                  (B) LOCATION 1...1311
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:119
       ACCACAAATA GAAAACCAAA TACTAATATG AAACTTTCAT CTAAGAAAAT CTTAGCAATC
45
       ATTGCATTGC TGACGATGGG ACATGCTGTG CAGGCACAGT TTGTTCCGGC TCCCACCACA
                                                                                 120
       GGGATTCGCA TGTCTGTCAC TACAACCAAG GCCGTAGGCG AAAAAATCGA ATTGTTGGTT
                                                                                 180
       CATTCCATAG AGAAGAAAGG CATCTGGATC GATCTCAATG GGGATGCCAC TTACCAACAA
                                                                                 240
       GGAGAGGAAA TAACCGTATT CGATGAGGCA TACCACGAAT ACACGATCGG GACGCAAACC
                                                                                 300
       CTCACTATCT ATGGTAATAC GACCCGATTG GGCTGTCGAT CTACCGGTGC AACGGCTGTC
                                                                                 360
50
       GATGTAACGA AAAACCCTAA TCTGACCTAT CTCGCATGCC CGAAAAATAA TCTGAAATCA
                                                                                 420
       TTGGACTTGA CGCAAAACCC AAAGCTGCTG CGAGTTTGGT GCGACTCTAA CGAAATAGAA
                                                                                 480
       AGTTTGGACC TGAGTGGCAA TCCGGCTTTG ATCATCCTCG GCTGTGACAG GAATAAGCTG
                                                                                 540
       ACTGAGCTGA AGACCGATAA CAACCCCAAG TTGGCCTCTC TTTGGTGTTC TGATAATAAC
                                                                                 600
       CTGACGGAGT TGGAACTCAG TGCCAATCCT CGTCTCAATG ATCTTTGGTG CTTCGGTAAT
                                                                                 660
55
       CGGATCACGA AACTCGATCT GAGTGCCAAT CCTCTATTGG TAACACTTTG GTGCAGTGAC
                                                                                 720
       AATGAGCTTT CGACCTTGGA TCTTTCCAAG AATTCGGACG TTGCTTACCT TTGGTGTTCA
                                                                                 780
       TCGAACAAAC TTACATCCTT GAATCTGTCG GGGGTGAAGG GACTGAGTGT TTTGGTTTGT
       CATTCCAATC AGATCGCAGG TGAAGAAATG ACGAAAGTGG TGAATGCTTT GCCCACACTA
       TCTCCCGGCG CAGGCGCTCA GAGCAAGTTC GTCGTTGTAG ACCTCAAGGA CACTGATGAG
60
       AAGAATATCT GTACCGTAAA GGATGTGGAA AAAGCTAAAA GTAAGAACTG GCGAGTATTT
                                                                                1020
       GACTTCAACG GTGATTCTGA CAATATGCTT CCATACGAAG GAAGTCCGAC ATCGAACTTG
                                                                                1080
       GCAGTAGATG CTCCCACTGT CAGGATATAT CCCAATCCGG TAGGAAGATA TGCGCTCGTC
                                                                                1140
      GAGATCCCCG AGTCTCTTTT AGGGCAGGAA GCTGCTTTAT ACGATATGAA TGGGGTAAAA
                                                                                1200
      GTCTATAGTT TCGCGGTAGA GTCTCTTCGT CAGAACATTG ACCTGACACA TCTTCCCGAC
                                                                                1260
65
      GGCACTTATT TCTTCCGTCT CGATAACTAT ACCACTAAGC TCATCAAACA G
                                                                                1311
       (2) INFORMATION FOR SEQ ID NO:120
70
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 954 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
75
```

WO 99/29870 PCT/AU98/01023

```
(ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
  5
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 10
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...954
            (xi) SEQUENCE DESCRIPTION: SEO ID NO:120
15
       AATCATTTGA AGACAAATAT CAAGATGAGA AAAACAATAA TTTTCTGCTT GTTGCTCGCC
       CTATTTGGCT GTTCTTGGGC ACAAGAAAGA GTCGATGAAA AAGTATTCTC CGCAGGAACA
                                                                                   120
       AGTATTTTTA GGGGCATCCT TGAAAAGGTG AAAGCACCGC TTATGTATGG AGATCGTGAG
GTATGGGGTA TGGCTCGTGC GAGCGAGGAT TTCTTTTTTA TACTTCCCGT TACGGATGAC
                                                                                   180
                                                                                   240
20
       CTCACTCCCG TGCTTTTCTA TAACCGTCTT ACAAACGAAC CCTGCTTTGT GTCAGACCAA
                                                                                   300
       GGAATAACTG AGTATTTCAA ATTCGCTCAA GAAGGTGATT ACATTGAAGT CGAAGGAAGC
                                                                                   360
       TCTGTATTCA TGGCGAATCT TTTGTACTAT CGTTTTTTCC CGACAAGAAT TACCTCCTAT
                                                                                   420
       AATGCTCCCA TTGAAGGTGT TGTGAGCAAG ACGGGAAATC CTGCTTTTAC AATCCCGATG
                                                                                   480
       CTCCCGGGGG TTTCTGATTG CATAGAAATC TCAAACAACC GCAAAGTCTT TCTGACCAAT
                                                                                   540
25
       CAATTAGGGG TTGTAAACAT CACTGACGGG ATGGAACCTC CGATTATTGC CGGAGTCTCT
                                                                                   600
       GCTTCCTATG GATCTTCCGT CCGGGTGTAT GGTCATGTCT CACAGCGGTG GGACATCATA
                                                                                   660
       GGCCATTGCT ATTTGGATAT CTACCCAACC AATTGCTATC CGCTCAGCAC GAAACCCGTT
GCAGGAGACG ATGAGGTTTT TGTCAAACAA CAAGGCAGGC AAATAGAGAT CGATAGCAAC
                                                                                   720
                                                                                   780
       AGCCCCATAG TCCAAGTGGT CGTATACGAT CTTGAGGGGA AAAGTGTTTT TCGCAAAAGA
                                                                                   840
30
       ATGACCGAAA ACGCTTATAC CCTATCCTTT AGAGCACCCA TGCTCGGCTT TATGACCATC
                                                                                   900
       ATGATCGAAA CACAAAATTC GATTATCAAT AAAAAACTTA ATGTTACACA GCTA
                                                                                   954
       (2) INFORMATION FOR SEQ ID NO:121
35
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1383 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRAHDEDHESS: double
40
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
45
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
50
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1383
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:121
       AGACGGGCAG TTAAAATCCG ATCACCTCCC CATATCCACT CATTATTTGT ACGTAAATGT
                                                                                    60
       CTTTTTTCAG ACTTTAAATA CTTATATTTA TCCCGAAAAA TTACCCAAGA GAGATTGGGT
                                                                                   120
       CGATTGTCGA TCCGTCTTAA GTCATACAAT CCAATAATCA TTATTGAAAT GAAAAAAACA
                                                                                   180
60
      ACCATTATTT CTTTGATTGT CTTCGGTGCT TTCTTTGCAG CCGTGGGCCA AACCAAGGAC
                                                                                   240
      AATTCTTCTT ACAAACCTTT TTCGAAAGAA GATATTGCCG GAGGAGTTTA CTCTCTCCCG
                                                                                   300
      ACTCAAAATC GTGCGCAGAA GGACAATGCC GAGTGGCTTC TTACAGCGAC CGTCTCCACA
                                                                                   360
      AACCAGTCTG CAGATACTCA CTTTATCTTC GATGAGAACA ACCGCTATAT CGCTCGTGAC
                                                                                   420
      ATAAAAGCCA ATGGGGTAAG AAAATCCACG GACTCCATTT ACTACGATGC CAACGGGCGA
                                                                                   480
65
      ATATCGCATG TGGATCTTTA TATCTCGTTC AGTGGCGGAG AGCCTGCACT CGACACCCGA
                                                                                   540
      TTCAAGTACA CCTATGATGA CGAGGGAAAG ATGACCGTGA GGGAAGTATT CATGCTGGTA
                                                                                   600
      ATGGATCCGA ATACACCTAT CTCACGCTTG GAATATCATT ATGATGCACA GGGCAGACTG
                                                                                   660
      ACCCACTGGA TTTCTTTTGC TTTCGGGGCA GAATCCCAAA AGAATACGTA TCACTATAAT
                                                                                   720
      GAAAAAGGTC TGTTGGTCAG CGAAGTGCTG AGCAATGCAA TGGGGACAAC CTATTCAGAC
                                                                                   780
70
      ACCGGCAAAA CGGAATACAG CTATGACGAT GCAGATAATA TGGTGAAGGC CGAGTACTTC GTCGTCCAGC AAGGAAAGGC ATGGCAAGTA CTCAAAAGAG AGGAATACAC CTATGAGGAC
                                                                                  840
                                                                                  900
      AATATCTGCA TACAATATTT GGCTATTAAC GGTACCGACA CAAAGGTGTA CAAGCGAGAC
                                                                                  960
      ATCGAGAGCG ATAAGTCCAT CTCCGCAAAT GTCATTGACA TTCCGTCAAT GCCGGAACAG
                                                                                 1020
      ACCTGGCCTA ATATGTACGG ATTCAACGCA AAGCGACTGA AAGAGACTTA TTCCTCCTAC
                                                                                 1080
75
      GAAGGAGATG TGGCTACTCC TATATTCGAC TATATCTATA CGTACAAGGC TCTTACCTCA
                                                                                 1140
```

WO 99/29870

| 5 | ATGGCAACAC CTTCGACAGA AGCTCAGGTA GCAGTCTATC TCAATCCGTC AACGGACCGG TTAGTGATTC TGGCCAACGG CATCACACAT CTGAGCATGT ACGACTTGCA GGGTAAGCTT ATCCGTGATT GTGCCTTGAG CGGCGATAAG GTGGAAATGG GTGTCGGATC TTTGACCAAA GGGACATACC TGCTTAAAGT GAATACGGAT CAGGGAGCCT TTGTGAGAAA AGTCGTGATT CGA | 1200 1260 1320 1380 1383 |
|------------|--|--|
| | (2) INFORMATION FOR SEQ ID NO:122 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1353 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double | |
| 15 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 20 | | |
| 20 | (iv) AHTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 25 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11353</pre> | |
| 30 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:122 | |
| 30 | ATGGCAAAAG TTATAAAAAC AAAAAAAGGC CTTGCACTTA ATCTGAAAGG AAAACCGCTG CCCGAGATGC TGGCCGAACC GGCCCAAAGT CCTACTTACG CGGTCGTGCC CGACGATTTT GAAGGTGTTTA TCCCCAAGGT GACGGCTCGT CCGGGGGATA AGGTGCGTGC CGGCTCAGCA | 60 120 180 |
| 35 | CTGATGCACC ACAAGGCATA TCCGGAGATG AAGTTTACAA GTCCGGTTAG CGCGAAAGTG ATCGCGGTGA ATCGCGGTGC CAAGCGCAAG GTGTTGAGCA TCGAGGTGAA ACCGGACGGA CTGAACGAAT ACGAGTCATT CCCTGTCGGG GATCCGTCTG CCCTCTCTGC CGAACAGATC AAGGAGCTTT TACTGTCGAG CGGTATGTGG GGTTTTATTA AGCAACGTCC TTACGACATA GTGGCTACAC CGGATATAGC TCCACGCGAC ATTTATATTA CTGCCAACTT TACTGCACCA | 240 300 360 420 |
| 40 | TTGGCTCUGG ACTTCGATTT CATCGTTCGA GGAGAAGAAC GCGCCCTGCA GACTGCCATC GATGCCTTGG CCAAACTCAC GACAGGAAAG GTGTATGTGG GCCTGAAGCC GGGTTCATCT CTGGGCTTGC ACAATGCAGA AATCGTAGAA GTACACGGAC CTCATCCGGC AGGTAACGTG GGCGTGCTGA TCAATCATAC GAAGCCAATC AATCGGGGCC AAACGTGTG GACGCTCAAG | 480 540 600 660 720 |
| 45 | GCTACCGACC TGATCGTGAT CGGACGTTTC CTGCTTACGG GCAAAGCCGA TTTTACCAGA ATGATTGCCA TGACCGGCTC AGACGCTGCA GCTCACGGAT ACGTCCGTAT TATGCCGGGT TGCAATGTCT TTGCTTCCTT CCCCGGCCGA CTGACAATAA AGGAATCTCA CGAGCGTGTG ATCGCATGACA ATGATCACGGT GATCCCCGAA GGCGACGATG TGGACGAACT CTTGCGGGGG GCTGCACCCC GTCTCAGCC GATCCCGAA GGCGACGATT TTTCTCTTT CTTCGGTGG GCTGCACCCC GTCTCAGTCA GTACCACGAC TATCTCTTT CTTCAGCC GCTGCACCCC GTCTCAGTCA GTACCACGAT AGCAGACTT ATTTCTCTTT GTTGCAGGGG | 780 840 900 960 1020 1080 |
| 50 | AAAAACAAAG AGTACGTACT CGATGCCCGG ATCAAGGGTG GCGAACGTGC TATGATCATG AGCAACGAGT ATGACCGCGT TTTCCCGATG GACATCTATC CGGAGTATTT GCTCAAGGCT ATTATAGCAT TCGACATCGA CAAGATGGAG GACTTTAGCA TATATGAAGT GGCTCCGGAG GACTTTGCCA CTTGCGAATT TGTGGATACA TCCAAGATCG AGCTGCAGCG TATCGTTCGC GAGGGCTTGG ATATGCTCTA TAAGGAAATG AAT | 1140 1200 1260 1320 1353 |
| 55 | | |
| | (2) INFORMATION FOR SEQ ID NO:123 | |
| 60 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 585 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 65 | (ii) MOLECULE TYPE: DNA (genomic) | |
| 00 | (iii) HTPOTHETICAL: NO | |
| | (iv) AUTI-SENSE: NO | |
| 70 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHTROHONAS GINGIVALIS | |
| <i>7</i> 5 | (ix) FEATURE: (A) NAHE/KEY: misc_feature (B) LOCATION 1585 | |

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123
      ATGGATAAAG TGAGCTATGC TCTGGGATTG AGCATCGGTA ATAATTTCAA GTCTTCGGGC
                                                                                   60
 5
       ATCGACAGCG TCGTTATGGA TGATTTCATG CAAGGTCTGT CTGATGTACT GGAAGAAAAA
                                                                                  120
       GCCCCTCAGC TCTCGTATGA CGAGGCCAAG CGCGAAATAG AGGCGTATTT CATGGATTTG
                                                                                  180
       CAGCAGAAGG CTGTCAAACT GAACAAAGAG GCCGGAGAAG AATTCCTCAA GATAAATGCA
                                                                                   240
      CACAAGGAAG GTGTGACGAC CTTACCGAGC GGCTTGCAAT ACGAAGTCAT TAAGATGGGA
GAGGGCCCGA AACCCACCCT TTCGGACACG GTAACCTGTC ATTATCACGG TACGCTCATC
                                                                                   300
                                                                                  360
10
      AACGGTATCG TTTTCGATAG CTCTATGGAC AGGGGAGAAC CGGCCAGTTT CCCTCTAAGA
GGAGTTATAG CCGGCTGGAC GGAGATTCTT CAATTAATGC CTGTAGGATC CAAGTGGAAA
                                                                                  420
                                                                                   480
      GTAACTATAC CGAGCGATCT GGCGTATGGA GATCGTGGTG CCGGCGAACA TATCAAACCG
                                                                                  540
      GGTAGTACGC TCATTTTTAT AATCGAATTA TTGAGTATCA ACAAA
                                                                                  585
15
      (2) INFORMATION FOR SEQ ID NO:124
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 819 base pairs
20
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
25
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SEHSE: NO
30
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
35
                 (B) LOCATION 1...819
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:124
      ATGAAAAAAG CATTACTTAT TGGTGCTGCT CTTTTGGGAG CAGTCAGTTT TGCAAGTGCT
40
      CAGTCTTTGA GCACAATCAA AGTACAGAAC AATTCAGTAC AGCAACCTCG TGAGGAAGCC
                                                                                  120
      ACTATTCAGG TTTGTGGAGA ATTGGCAGAG CAAGTTGACT GCATTGGGAC AGGTAATTCT
                                                                                  180
      GCAATCATAG CCGCTGCAGC GAAATTTGAA AGCGATGATC TCGAAAGCTA TGTTGGCTGG
      GAGATCATGA GTGTTGATTT CTTCCCTGGA TATAAAGCGT GCAAGTACAC ATCTGCAGTC
                                                                                  300
      TGGGCTGATG ATATGACCAT TTTGGGCCAA TCAGAAGATA GTGATCCCGA AATGCAGACT
                                                                                  360
45
      ATCAACAATC TTGCTCTCAA GACTAGTGTC AAGATTGAAG CCGGCAAGAA TTACATAGTT
                                                                                  420
      GGTTATATTG CTAATACCGC AGGTGGACAT CCTATCGGAT GTGATCAGGG CCCTGCCGTT
                                                                                  480
      GATGGTTATG GAGATTTGGT TTCTATATCA GAAGATGGTG GTGCTACTTT CCCTCCGTTC
                                                                                  540
      GAATCTCTTC ATCAAGCAGT TCCTACCTTA AATTACAACA TCTATGTCGT TGTTCATTTG
                                                                                  600
      AAGAAGGGTG AAGGTGTTGA GGCTGTTCTT ACCAACGACA AGGCTAATGC TTATGTTCAG
                                                                                  660
50
      AATGGCGTTA TCTATGTAGC CGGAGCTAAT GGTCGTCAGG TATCTCTGTT CGACATGAAC
                                                                                  720
      GGTAAGGTTG TTTATACCGG CGTTAGCGAA ACGATTGCAG CTCCTCAGAA GGGCATGTAT
                                                                                  780
      ATCCTCCGTG TAGGTGCTAA GAGCATCAAG CTGGCTATC
                                                                                  819
55
      (2) INFORMATION FOR SEQ ID NO:125
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1662 base pairs
                 (B) TYPE: nucleic acid
60
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
65
          (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
70
                 (A) ORGANISM: PORYPHYROMOMAS GINGIVALIS
           (ix) FEATURE:
                 (A) MAME/KEY: misc_feature
                 (B) LOCATION 1...1662
75
```

| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:125 | |
|------------|---|--------------------------------------|
| 5 | ATGCCAAGAA TTATGAAATT AAAAATTGCA CTCAGACTGC TGCTGGCGAC TTTTGCCATA GTTTTATTTA GCCCTCTGGC CAAGGCCCAG ATGGATATTG GTGGAGACGA TGTATTGATC GAGACGATGT CCACCCTATC AGGATATTCA GAGGATTTTT ATTACAAGAT GGCTGTGGCA GACAATGGAT GGATCTATGT GATGTTGGAT TTCTCTCGTA TTTATTTTGA TGATGTCAGG CTGTATCGTT CCAAAGACGG TGGTGCTACT TACCAAAAGT TAGGGTCTTT GGGGTCTTTG | 60 120 180 240 300 |
| 10 | GTGCCTTATG ACTTCGATGT CTCGCATTGC GATTTTATTG TAACGGGAAA GGATGAAGAT GATATCAATG TTTGGACAGT CATGACAGCA TTCGAATATG TAGGTGGTAC TATTGGCAAT GGCGTTTTGC TGATGCATCG CCATGATGCA GATATCAATA ATACAGAGTG TGTGTACAAG | 360 420 480 |
| | AAGGATTTCC CTAATAATAG ACTGATGGGT GTAGCCATCG CCTCCAACTA CCGTGCGCCC TCTCCTTACG GTTTGGGGGG CGATCCTTTT GCTCTCGCTG TCGCCGTTAG TGGCTCCGGA AGCGATCACA GCTTCTTGGA CTATATTTTT TCGTTAGATG GTGGAGTACA CTTTGAGCAA | 540 600 660 |
| 15 | AGGGTATTT ACACAAGACC CCAAAACTG ACTATCAATA GAGTAGACCT TTCATTAGGC AGTACATCTC CTTCTCTTGG ATTTAATACT TGGCCACTAA TGGGAGTCGT ATTCGAAATG AATAAGAACC TTGATGGCTT CGACATTGGT TTCATTTCCA ACTTTTGTGGA CTATGATCCC CGCTATGCGT GGTCTGAACC GATAATAATA GAAGAAGACT GTGCATGGAC TGATTTTAAT CCTTTGGGAG CACTAAGTAT AGAGATCCAA ATGATGTTGG ATGACAATTC GGATAATAACC | 720 780 840 900 960 |
| 20 | GTGGGTGGAG AACGCTCCCA TAACTTCCTG ATCACTTACC CGGGCCATTA CGTATATCCG AAGCAATCTT TCAATTATTC TCCCGGACAT ACACCGACAA AGAAAGATCT GGTCTTTAAA CACTGTATAG GTATTCCGGC TTTGGCCATAC GATAAGGAAG GCGATCGTTA TCTGACTACT TTTCAAGGATC ACAATCTAAT GAGATACAGA TGGATCAAAT ACGATGACAT TAACTCTTTT TATGGTTGGA GTTGGCCATA TGTATATCCA AAAGAAGCTA AAGATAAAAA GAGGCGCCGT | 1020 1080 1140 1200 1260 |
| 25 | CCGCAAGTAG CACTCAATCC TACCAATGGA AAGGCTTGTT GGGTATGGCA TACTCGCAAG AGCCCATATG ATGAAACCAA ACCACATCCT ACTCCTGTAA TTATTAAACA TTTCCTATGG TCCGATACGG AGTGGGTACA TGCTCTGGAC GTGGGGGGAC TATTGCAGAA GGAGGGTAGC ATGAAGCTCT ACCCCAATCC TGCCAAAGAA TATGTTCTGA TCAACCTACC CAAAGAAGGG GGGCACGAGG CAGTCGTATA CGACATGCAG GGCCGAATCG TGGAGAAAGT TTCATTTTCA | 1320 1380 1440 1500 1560 |
| 30 | GGGAAAGAAT ATAAGCTGAA TGTGCAGTAT CTGTCCAAAG GTACGTACAT GCTGAAAGTT GTAGCGGATA CGGAGTATTT CGTGGAAAAA ATCATTGTAG AG | 1620 1662 |
| | (2) INFORMATION FOR SEQ ID NO:126 | |
| 35 | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1650 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double | |
| 40 | (D) TOPOLOGY: circular (ii) MOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 45 | (iv) ANTI-SENSE: NO (vi) ORIGINAL SOURCE: | |
| 50 | (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 50 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11650</pre> | |
| 55 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:126 ATGAAATTAA AAATTGCACT CAGACTGCTG CTGGCGACTT TTGCCATAGT TTTATTTAGC | 60 |
| 60 | CCTCTGGCCA AGGCCCAGAT GGATATTGGT GGAGACGATG TATTGATCGA GACGATGTCC ACCCTATCAG GATATTCAGA GGATTTTTAT TACAAGATGG CTGTGGCAGA CAATGGATGG ATCTATGTGA TGTTGGATTT CTCTCGTATT TATTTTGATG ATGTCAGGCT GTATCGTTCC AAAGACGGTG GTGCTACTTA CCAAAAGTTA GGGTCTTTGG GGTCTTTGGT GCCTTATGAC | 120 180 240 300 |
| | TTCGATGTCT CGCATTGCGA TTTTATTGTA ACGGGAAAGG ATGAAGATGA TATCAATGTT TGGACAGTCA TGACAGCATT CGAATATGTA GGTGGTACTA TTGGCAATGG CGTTTTGCTG ATGCATCGCC ATGATGCAGA TATCAATAAT ACAGAGTGTG TGTACAAGAA GGATTTCCCT AATAATAGAC TGATGGGTGT AGCCATCGCC TCCAACTACC GTGCGCCCTC TCCTTACGGT | 360 420 480 540 |
| 65 | TTGGGGGGCG ATCCTTTGC TCTCGCTGTC GCCGTTAGTG GCTCCGGAAG CGATCACAGC TTCTTGGACT ATATTTTTTC GTTAGATGGT GGAGTACACT TTGAGCAAAA GCGTATTTAC ACAAGACCCC AAAAACTGAC TATCAATAGA GTAGACCTT CATTAGGCAG TACATCTCT TCTCTTGGAT TTAATACTTG GCCACTAATG GGAGTCGTAT TCGAAATGAA TAAGAACCTT | 600 660 720 780 |
| 70 | GATGGCTTCG ACATTGGTTT CATTTCCAAC TTTGTGGACT ATGATCCCCG CTATGCGTGG TCTGAACCGA TAATAATAGA AGAAGACTGT GGATGGACTG ATTTTAATCC TTTGGGAGCA CTAAGTATAG AGATCCAAAT GATGTTGGAT GACAATTCGG ATAATACCGT GGGTGGAGAA CGCTCCCATA ACTTCCTGAT CACTTACCCG GGCCATTACG TATATCCGAA GCAATCTTTC | 840 900 960 1020 |
| <i>7</i> 5 | AATTATTCTC CCGGACATAC ACCGACAAAG AAAGATCTGG TCTTTAAACA CTGTATAGGT ATTCCGGCTT TGGCATACGA TAAGGAAGGC GATCGTTATC TGACTACTTT TCAAGATCAC AATCTAATGA GATACAGATG GATCAAATAC GATGACATTA ACTCTTTTTA TGGTTGGAGT | 1080 1140 1200 |

WO 99/29870 PCT/AU98/01023

```
TGGCCATATG TATATGCAAA AGAAGCTAAA GATAAAAAGA GGCGCCGTCC GCAAGTAGCA
                                                                                1260
      CTCAATCCTA CCAATGGAAA GGCTTGTTGG GTATGGCATA CTCGCAAGAG CCCATATGAT
                                                                                1320
      GAAACCAAAC CACATCCTAC TCCTGTAATT ATTAAACATT TCCTATGGTC CGATACGGAG
                                                                                1380
      TGGGTACATG CTCTGGACGT GGGGGACGTA TTGCAGAAGG AGGGTAGCAT GAAGCTCTAC
                                                                                1440
 5
      CCCAATCCTG CCAAAGAATA TGTTCTGATC AACCTACCCA AAGAAGGGGG GCACGAGGCA
                                                                                1500
      GTCGTATACG ACATGCAGGG CCGAATCGTG GAGAAAGTTT CATTTTCAGG GAAAGAATAT
                                                                                1560
      AAGCTGAATG TGCAGTATCT GTCCAAAGGT ACGTACATGC TGAAAGTTGT AGCGGATACG
                                                                                1620
      GAGTATTTCG TGGAAAAAAT CATTGTAGAG
                                                                                1650
10
      (2) INFORMATION FOR SEQ ID NO:127
            (i) SEQUENCE CHARACTERISTICS:
15
                 (A) LENGTH: 1170 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
20
           (ii) HOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
25
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
30
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1170
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:127
35
      ATGAAACGAT TACTCCCCTT TCTCCTTTTA GCAGGACTCG TAGCCGTAGG AAACGTGTCT GCTCAGTCAC CCCGAATCCC TCAAGTGGAT GTACACACTC GCATCGCAAG AAATGCCCGT
                                                                                  60
                                                                                 120
      TATCGACTGG ACAAGATCAG TGTCCCGGAT TCTCGTCAGA TATTCGATTA CTTCTATAAA
                                                                                 180
      GAAGAAACGA TACCCACTAA AATACAAACG ACCACAGGAG GTGCAATTAC AAGCATCGAT
                                                                                 240
      TCGCTTTTCT ATGAAGACGA CAGGTTGGTT CAGGTGCGCT ATTTTGACAA TAACCTTGAA
                                                                                 300
40
      TTAAAACAAG CGGAGAAGTA TGTATACGAC GGTTCTAAGC TGGTCCTTCG AGAAATTCGC
                                                                                 360
      AAGTCGCCGA CAGACGAAAC GCCAATAAAG AAAGTTAGCT ATCACTATCT CTGTGGCAGC
                                                                                 420
      GATATGCCTT TTGAGATTAC GACAGAGATG AGCGATGGCT ATTTTGAAAG CCATACGCTT
                                                                                  480
      AACTATCTGA ATGGAAAGAT TGCCCGAATA GATATCATGA CTCAACAGAA CCCATCGGCC
                                                                                 540
      GAATTGATCG AAACGGGTAG AATGGTATAT GAGTTTGATG CCAATAATGA TGCTGTACTG
                                                                                 600
45
      CTTCGTGACA GTGTATTTCT TCCTCTTCAA AACAAGTGGG TAGAAATGTT TACTCACCGT
                                                                                 660
      TATACATACG ACAATAAGCA TAATTGTATT CGTTGGGAAC AAGACGAATT CGGCACCCTC
                                                                                 720
      ACCCTTGCCA ACAACTTCGA ATACGACACC ACTATCCCTC TGTCGTCTGT ATTGTTCCCC
                                                                                 780
      ACGCATGAGG AGTTCTTCCG TCCTCTTCTT CCCAATTTTA TGAAGCATAT GCGTACGAAG CAAACGTATT TCAATAACTC CGGAGAAGGC TTGTCAGAGG TATGCGATTA CAACTACTTC
                                                                                 840
                                                                                 900
50
      TATACCGATA TGCAGGGTAA TGCACTGACC GATGTTGCCG TGAACGAATC GATCAAGATT
                                                                                 960
      TATCCTCGTC CTGCCACGGA TTTTCTGCGT ATAGAAGGTT CGCAACTGCT TCGCCTTTCG
                                                                                1020
      CTATTCGACA TGAACGGGAA GCTCATCAGA GCTACCGAAT TGACAGGCGA TTTGGCCATT
                                                                                1080
      ATCGGAGTTG CATCTCTCC GAGAGGCACT TACATCGCAG AAATAACTGC TGCAAACAGC
                                                                                1140
      AAAACCATAC GTGCAAAAGT ATCGCTCAGA
                                                                                1170
55
      (2) INFORMATION FOR SEQ ID NO:128
            (i) SEQUENCE CHARACTERISTICS:
60
                 (A) LENGTH: 1233 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
65
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) AUTI-SENSE: NO
70
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROHONAS GINGIVALIS
           (ix) FEATURE:
75
                 (A) NAME/KEY: misc feature
```

```
(B) LOCATION 1...1233
```

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128
```

| Э | ATGAGACAGC | ATTTATCTCT | ATTTCCTTTT | ATCTTGTTTC | TGCTTCTTGC | CTTCTCTTAT | 60 |
|----|------------|------------|------------|------------|------------|------------|------|
| | GTCGGTTGCA | GAACAGTCCG | ACAAACACCT | AAGCAGTCGG | AACGGTACGT | CGTAGTCCTG | 120 |
| | TCTTTGGACG | GCTTCCGACC | GGACTATACC | GATCGGGCAC | GTACACCGGC | GTTGGATCGG | 180 |
| | ATGGCACAGG | AGGGATTGAG | CGGGTCGCTC | CAACCATGCT | TCCCCTCGCT | TACATTTCCC | 240 |
| | AATCATTACA | GCATGGCTAC | GGGGCTTTAC | CCCGATCATC | ACGGTATCGT | AGCCAATGAG | 300 |
| 10 | TTTGTGGATT | CGCTACTGGG | CATCTTTCGT | ATATCCGACC | GAAAAGCCGT | GGAGACCCCC | 360 |
| | GGATTTTGGG | GCGGCGAGCC | GGTTTGGAAT | ACGGCCGCAC | GCCAAGGCAT | CCGTACCGGT | 420 |
| | GTCTACTTTT | GGGTAGGATC | CGAAACGGCT | GTGAACGGAA | ATCGGCCGTG | GCGGTGGAAA | 480 |
| | AAATTCTCCT | CCACCGTTCC | GTTTCGTGAC | CGTGCCGACT | CCGTCATCGC | GTGGCTCGGA | 540 |
| | CTGCCCGAAA | AGGAGCGACC | GCGCTTGCTC | ATGTGGTACA | TCGAGGAGCC | GGATATGATC | 600 |
| 15 | GGACACAGCC | AAACGCCCGA | AAGCCCGCTG | ACACTGGCAA | TGGTAGAGCG | GTTGGACAGT | 660 |
| | GTGGTCGGCT | ATTTCCGCAA | GCGGTTGGAC | TCTCTGCCCA | TAGCCGCACA | GACCGACTTC | 720 |
| | ATCATAGTAT | CCGATCACGG | TATGGCCACG | TACGAAAATG | AGAAATGTGT | CAATCTGTCG | 78C |
| | CATTATCTGC | CTGCGGACAG | TTTCCTCTAC | ATGGCCACCG | GGGCCTTCAC | CCACTTGTAC | 840 |
| | CCGAAGCCCT | CCTATACCGA | GCGAGCCTAT | GAGATCCTGC | GGGCCATTCC | ACATATATCG | 900 |
| 20 | GTTTACCGCA | AGGGGGAGGT | GCCCAAGCGT | TTGCGCTGTG | GCACCAATCC | TCGTTTGGGC | 960 |
| | GAACTGGTCG | TGATTCCGGA | CATAGGCTCC | ACCGTCTTTT | TCGCAATAAA | TGAAGACGTT | 1020 |
| | CGTCCGGGAG | CGGCACATGG | CTATGACAAC | CAAGCACCGG | AAATGCGGGC | TTTACTCCGG | 1080 |
| | GCTGTCGGAC | CCGATTTCCG | TCCGGGCAGT | AGGGTGGAAA | ACCTGCCGAA | TATCACCATC | 1140 |
| ~- | TATCCGCTCA | TATGCAGGCT | GTTGGGTATA | GAGCCTGCAC | CCAACGATGC | GGACGAAACG | 1200 |
| 25 | TTGCTGAACG | GCCTGATCCG | AGACAAACGA | CCA | | | 1233 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO:129

| 30 | (i) | SECUENCE | CHARACTERISTICS: |
|----|-------|-----------|------------------|
| 30 | 111 | SECOPINCE | CHARACTERISTICS: |

- (A) LENGTH: 738 base pairs
- (B) TYPE: nucleic acid (C) STRAMDEDNESS: double
- (D) TOPOLOGY: circular 35

(ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO
- 40 (iv) ANTI-SENSE: NO
 - (vi) ORIGINAL SOURCE:

(A) ORGANISM: PORYPHYROMONAS GINGIVALIS

- 45 (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...738

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129 50

| | ATGAAAGTAG | GTTTGTTCAT | CCCCTGTTAT | GTCAATGCAG | TGTATCCGGA | AGTGGGTATC | 60 |
|----|------------|------------|------------|------------|------------|------------|-----|
| | GCCACGTACA | AACTGCTGAA | GAGTTTGGAC | ATAGATGTCG | ACTACCCGAT | GGATCAGACA | 120 |
| | TGTTGCGGCC | AGCCTATGGC | CAATGCCGGA | TTCGAACAGA | AAGCTCAAAA | GCTGGCTTTG | 180 |
| | CGATTCGAAG | AGCTGTTCGA | GTCGTATGAT | GTAGTCGTAG | GGCCATCGGC | CAGTTGCGTT | 240 |
| 55 | GCTTTCGTGA | AAGAAAACTA | TGATCATATC | CTCAGACCGA | CAGGACATGT | CTGCAAGTCG | 300 |
| | GCAGCCAAGG | TTCGGGATAT | ATGCGAGTTC | TTGCACGATG | ACCTGAAGAT | CACCAGCCTC | 360 |
| | CCCTCCCGAT | TCGCCCATAA | GGTGAGCCTG | CACAACAGTT | GCCACGGTGT | GCGCGAACTG | 420 |
| | CATCTGTCCA | CCCCCAGTGA | AGTGCACCGA | CCGTACCACA | ACAAGGTGCG | CCGGCTATTG | 480 |
| ~~ | GAGATGGTGC | AGGGCATAGA | GGTATTCGAG | CCGAAGCGAA | TAGACGAATG | CTGCGGTTTC | 540 |
| 60 | GGCGGTATGT | ACTCGGTGGA | GGAGCCGGAG | GTATCCACCT | GTATGGGGCA | TGACAAGGTG | 600 |
| | CTGGATCACA | TATCCACAGG | TGCGGAGTAC | ATCACAGGGC | CGGACAGCTC | GTGCCTCATG | 660 |
| | CATATGCAGG | GAGTGATAGA | CAGAGAGAAA | TTGCCGATCA | AGACAATTCA | TGCAGTAGAA | 720 |
| | ATTTTAGCAG | CAAACTTA | | | | | 738 |

65 (2) INFORMATION FOR SEQ ID NO:130

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 738 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic) 75

70

```
(iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
 5
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
10
                  (B) LOCATION 1...738
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:130
       ATGGATATIG TAAGTATGGC CGATAAAGCT CTTGTAGTGG AGATGAGAGA TGTGACGCTC
15
       TGTCAGGAGG AAAACGTCAT TTTTCAAAAT TTGAATCTGA CCCTTTCCGC CGGAGACTTC
       GTCTATCTGA TAGGCTCAGT GGGATCGGGG AAGAGCACTT TGCTGAAGGC TTTGTATGCT
                                                                                    180
       GAGGTGCCTA TCTCTGCCGG TTATGCCCGC GTGATAGATT ATGATCTGGC AAAGTTGAAA
                                                                                    240
       CGGAAGCAGT TGCCCTATCT GCGCAGGAAT TTGGGCATTG TGTTTCAGGA TTTCCAGTTG
                                                                                    300
       CTGAACGGAC GTACTGTTGC GGAGAATTTG GATTTCGTTT TGCGAGCTAC GGACTGGAAA
                                                                                    360
       AACCGAGCCG ATCGCGAGCA GCGTATCGAG GAGGTTTTGA CCCGTGTGGG AATGTCTCGG
AAGGCTTATA AGAGACCGCA CGAACTGTCC GGAGGGGAGC AACAACGTGT GGGTATAGCC
20
                                                                                    420
                                                                                    480
       AGAGCTTTGC TGGCGAAGCC TGCGTTGATC CTGGCCGACG AACCCACAGG CAACCTCGAT
                                                                                    540
       TCGGTGACCG GATTGCAGAT CGCTTCTCTG CTCTACGAAA TCAGTAAGCA GGGCACTGCA
                                                                                    600
       GTACTTATGA GCACGCACAA CAGCAGCCTG CTGTCGCATC TGCCGGCACG GACATTGGCC
                                                                                    660
25
       GTTCGTAAGA ATGGCGATGC CTCCTCTTTG GTCGAGCTGA GTGCAGATGC TGTTTCAAGA
                                                                                    720
       AAAAATACGG AAATAGAT
       (2) INFORMATION FOR SEQ ID NO:131
30
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 723 base pairs
                  (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
35
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
40
           (iv) AHTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
45
           (ix) FEATURE:
                  (A) NAME/KET: misc_feature
                  (B) LOCATION 1...723
50
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:131
       ATGGCCGATA AAGCTCTTGT AGTGGAGATG AGAGATGTGA CGCTCTGTCA GGAGGAAAAC
       GTCATTTTTC AAAATTTGAA TCTGACCCTT TCCGCCGGAG ACTTCGTCTA TCTGATAGGC
       TCAGTGGGAT CGGGGAAGAG CACTTTGCTG AAGGCTTTGT ATGCTGAGGT GCCTATCTCT
                                                                                    180
55
       GCCGGTTATG CCCGCGTGAT AGATTATGAT CTGGCAAAGT TGAAACGGAA GCAGTTGCCC
TATCTGCGCA GGAATTTGGG CATTGTGTTT CAGGATTTCC AGTTGCTGAA CGGACGTACT
                                                                                    240
                                                                                    300
       GTTGCGGAGA ATTTGGATTT CGTTTTGCGA GCTACGGACT GGAAAAACCG AGCCGATCGC
GAGCAGCGTA TCGAGGAGGT TTTGACCCGT GTGGGAATGT CTCGGAAGGC TTATAAGAGA
                                                                                    360
                                                                                    420
       CCGCACGAAC TGTCCGGAGG GGAGCAACAA CGTGTGGGTA TAGCCAGAGC TTTGCTGGCG
                                                                                    480
60
       AAGCCTGCGT TGATCCTGGC CGACGAACCC ACAGGCAACC TCGATTCGGT GACCGGATTG
                                                                                    540
       CAGATCGCTT CTCTGCTCTA CGAAATCAGT AAGCAGGGCA CTGCAGTACT TATGAGCACG
                                                                                    600
       CACAACAGCA GCCTGCTGTC GCATCTGCCG GCACGGACAT TGGCCGTTCG TAAGAATGGC
                                                                                    660
       GATGCCTCCT CTTTGGTCGA GCTGAGTGCA GATGCTGTTT CAAGAAAAA TACGGAAATA
                                                                                    720
       GAT
                                                                                    723
65
       (2) INFORMATION FOR SEQ ID NO:132
             (i) SEQUENCE CHARACTERISTICS:
70
                  (A) LEMGTH: 696 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRAHDEDHESS: double
                  (D) TOPOLOGY: circular
75
           (ii) HOLECULE TYPE: DNA (genomic)
```

```
(iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
 5
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROHONAS GINGIVALIS
           (ix) FEATURE:
10
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...696
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:132
15
      ATGAGAGATG TGACGCTCTG TCAGGAGGAA AACGTCATTT TTCAAAATTT GAATCTGACC
      CTTTCCGCCG GAGACTTCGT CTATCTGATA GGCTCAGTGG GATCGGGGAA GAGCACTTTG
                                                                              120
      CTGAAGGCTT TGTATGCTGA GGTGCCTATC TCTGCCGGTT ATGCCCGCGT GATAGATTAT
                                                                              180
      GATCTGGCAA AGTTGAAACG GAAGCAGTTG CCCTATCTGC GCAGGAATTT GGGCATTGTG
                                                                              240
      TTTCAGGATT TCCAGTTGCT GAACGGACGT ACTGTTGCGG AGAATTTGGA TTTCGTTTTG
                                                                              300
20
      CGAGCTACGG ACTGGAAAAA CCGAGCCGAT CGCGAGCAGC GTATCGAGGA GGTTTTGACC
                                                                              360
      CGTGTGGGAA TGTCTCGGAA GGCTTATAAG AGACCGCACG AACTGTCCGG AGGGGAGCAA
                                                                              420
      CAACGTGTGG GTATAGCCAG AGCTTTGCTG GCGAAGCCTG CGTTGATCCT GGCCGACGAA
                                                                              480
      CCCACAGGCA ACCTCGATTC GGTGACCGGA TTGCAGATCG CTTCTCTGCT CTACGAAATC
                                                                              540
      AGTAAGCAGG GCACTGCAGT ACTTATGAGC ACGCACAACA GCAGCCTGCT GTCGCATCTG
                                                                              600
25
      CCGGCACGGA CATTGGCCGT TCGTAAGAAT GGCGATGCCT CCTCTTTGGT CGAGCTGAGT
                                                                              660
      GCAGATGCTG TTTCAAGAAA AAATACGGAA ATAGAT
                                                                              696
      (2) INFORMATION FOR SEQ ID NO:133
30
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 657 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
35
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
40
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...6\overline{5}7
50
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:133
      ATGATTGAAA TCAGCAACCT CACCAAGGTT TTCAGAACAG AAGAAATAGA GACGGTAGCC
      CTCGATGGCG TATCGCTCAA AGTGGACAAA GGCGAATTTA TCGCCATAAT GGGGCCTTCG
                                                                              120
      GGATGCGGTA AGTCCACTCT GCTCAATATC CTCGGCCTTC TCGACAATCC CACTTCCGGT
                                                                              180
55
      ATCTACAAGC TCGATGGGGC AGAAGTGGGC AACCTCCGGG AAAAAGACAG GACTGCCGTC
                                                                              240
      CGTAAGGGCA ATATCGGCTT CGTATTCCAG AGCTTCAACC TCATCGAAGA GATGACGGTA
                                                                              300
      AGCGAGAACG TGGAGTTGCC GCTCGTCTAT CTGGGTGTGA AGGCCTCCGA GCGGAAAGAG
                                                                              360
      CGAGTGGAGG AGGCACTGCG CAAGATGAGC ATCAGCCACC GGGCCGGCCA CTTCCCCAAT
      CAGCTCTCCG GAGGACAACA GCAGCGCGTG GCTATCGCCC GTGCCGTGGT GGCCAATCCG
                                                                              480
60
      AAGCTCATCC TCGCCGATGA ACCCACGGGT AACCTCGACT CCAAAAACGG AGCCGATGTC
                                                                              540
      ATGGAACTGC TCAGAGGTCT CAATCGCGAA GGTGCAACCA TCGTCATGGT GACGCACTCC
                                                                              600
      GAGCACGATG CACGTAGTGC CGGCCGCATC ATCAATCTGT TCGACGGTAA GATTCGC
65
      (2) INFORMATION FOR SEQ ID NO:134
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1785 base pairs
                 (B) TYPE: nucleic acid
70
                (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
75
         (iii) HYPOTHETICAL: NO
```

```
(iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
 5
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1785
10
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:134
      ATGAAAGAAT TTTTCAAAAT GTTTTTCGCC TCGATCCTCG GGGTTATAAC GGCAGGAATC
      ATCTTGTTCT GTATCTTCT ATTTATCTTT TTCGGCATCG TAGCCGGTAT TGCCTCCAAG
                                                                                120
15
      GCAACGGGAG GAACCATTCC GAAGATCGAA GCAAACTCCA TCCTACATAT ANACAATTCT
                                                                                180
      TCTTTCCCTG AGATCGTATC GGCCAATCCC TGGAGCATGC TCACAGGCAA AGACGAGTCC
                                                                                 240
      GTATCGCTCT CACAGGCAGT CGAAGCCATC GGCCAAGCCA AAAATAATCC CAACATAACC
                                                                                 300
      GGTATCTTCC TCGATCTGGA CAACCTTTCC GTCGGTATGG CATCGGCAGA GGAATTGCGT
                                                                                360
      CGCGCGTTGC AGGATTTCAA GATGTCGGGC AAGTTCGTCG TATCCTATGC CGACAGATAC
                                                                                420
20
      ACCCAAAAGG GTTACTACCT CTCCAGTATT GCAGACAAAC TCTACCTCAA TCCGAAAGGA
                                                                                 480
      ATGTTGGGGC TTATCGGGAT TGCGACCCAA ACAATGTTCT ACAAAGATGC CCTCGACAAA
                                                                                 540
      TTCGGCGTGA AGATGGAGAT CTTCAAGGTA GGCACCTACA AGGCAGCCGT AGAGCCATTC
      ATGCTCAACA GGATGAGCGA TGCCAATCGC GAACAATCA CCACATACAT AAACGGGCTT
                                                                                 660
      TGGGACAAGA TCACATCCGA TATTGCAGAG TCGCGCAAGA CGGCAATGGA TTCCGTGAAA
                                                                                720
      ATGTTTGCCG ACAAAGGCGA AATGTTCGGT CTTGCCGAGA AAGCGGTGGA GATGAAGCTC
GTGGATGAGC TGGCTTACCG TACCGATGTG GAGAAAGAAC TCAAAAAGAT GTCCCAACGC
25
                                                                                 780
      GGAGAGAAAG ATGAACTTCG GTTCGTATCG CTTTCTCAGG TTCTGGCCAA TGGCCCGATG
                                                                                900
      AACAAAACGA AAGGCAGTCG GATCGCCGTT CTCTTTGCCG AAGGTGAAAT AACGGAAGAA
                                                                                960
      ATAATAAAGA AGCCGTTCGA CACTGACGGT AGCTCCATCA CACAAGAACT CGCCAAAGAA
                                                                               1020
30
      ATCAAGGCAG CAGCCGATGA CGATGATATC AAAGCCGTAG TACTTCGTGT CAATTCTCCG
                                                                               1080
      GGAGGTAGTG CTTTCACTTC CGAACAGATA TGGAAGCAGG TAGCCGATCT CAAGGCCAAA
                                                                               1140
      AAGCCTATCG TGGTCTCCAT GGGCGACGTA GCAGCCTCGG GCGGATACTA CATAGCCTGC
                                                                               1200
      GCAGCCAACA GTATCGTGGC AGAGCATACG ACTCTGACCG GCTCCATCGG CATATTCGGC
                                                                               1260
      ATGTTCCCGA ACTTCGCGGG CGTAGCCAAG AAGATAGGAG TGAATATGGA CGTCGTACAG
                                                                               1320
35
      ACATCCAAGT ATGCAGACTT GGGCAACACC TTCGCTCCGA TGACGGTCGA AGATCGTGCC
                                                                               1380
      CTCATCCAAC GCTACATAGA GCAGGGCTAC GACCTCTTCC TCACTCGCGT ATCGGAAGGC
                                                                               1440
      CGCAACCGCA CCAAGGCACA GATCGACAGC ATCGCTCAAG GCCGTGTATG GCTCGGCGAC
                                                                               1500
      AAAGCTCTTG CACTCGGTTT GGTGGATGAG CTTGGAGGTT TGGACACAGC TATCAAACGG
GCCGCGAAGC TGGCTCAGCT CGGTGGCAAC TACAGCATAG AGTATGGCAA GACCAAGCGC
                                                                               1560
                                                                               1620
40
      AACTTCTTCG AAGAGTTGCT CTCCTCATCA GCAGCGGATA TGAAGTCTGC CATCCTGAGT
                                                                               1680
      ACCATTCTCT CCGATCCGGA AATAGAAGTT CTGCGCGAAC TCCGCTCCAT GCCGCCCCGT
                                                                               1740
      CCTTCGGGCA TACAGGCACG TCTCCCCTAT TACTTCATGC CGTAC
                                                                                1785
45
      (2) INFORMATION FOR SEQ ID NO:135
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1767 base pairs
                 (B) TYPE: nucleic acid
50
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
55
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
60
                 (A) ORGANISM: PORYPHYROMOMAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1767
65
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:135
      ATGTTTTTCG CCTCGATCCT CGGGGTTATA ACGGCAGGAA TCATCTTGTT CTGTATCTTT
      CTATTTATCT TTTTCGGCAT CGTAGCCGGT ATTGCCTCCA AGGCAACGGG AGGAACCATT
                                                                                 120
70
      CCGAAGATCG AAGCAAACTC CATCCTACAT ATAHACAATT CTTCTTTCCC TGAGATCGTA
                                                                                 180
      TCGGCCAATC CCTGGAGCAT GCTCACAGGC AAAGACGAGT CCGTATCGCT CTCACAGGCA
                                                                                 240
      GTCGAAGCCA TCGGCCAAGC CAAAAATAAT CCCAACATAA CCGGTATCTT CCTCGATCTG
                                                                                 300
      GACAACCTTT CCGTCGGTAT GGCATCGGCA GAGGAATTGC GTCGCGCGTT GCAGGATTTC
                                                                                 360
      AAGATGTCGG GCAAGTTCGT CGTATCCTAT GCCGACAGAT ACACCCAAAA GGGTTACTAC
                                                                                 429
75
      CTCTCCAGTA TTGCAGACAA ACTCTACCTC AATCCGAAAG GAATGTTGGG GCTTATCGGG
                                                                                 480
```

```
ATTGCGACCC AAACAATGTT CTACAAAGAT GCCCTCGACA AATTCGGCGT GAAGATGGAG
                                                                            540
      ATCTTCAAGG TAGGCACCTA CAAGGCAGCC GTAGAGCCAT TCATGCTCAA CAGGATGAGC
                                                                            600
      GATGCCAATC GCGAACAAAT CACCACATAC ATAAACGGGC TTTGGGACAA GATCACATCC
                                                                            660
      GATATTGCAG AGTCGCGCAA GACGGCAATG GATTCCGTGA AAATGTTTGC CGACAAAGGC
                                                                            720
      GAAATGTTCG GTCTTGCCGA GAAAGCGGTG GAGATGAAGC TCGTGGATGA GCTGGCTTAC
                                                                            780
      CGTACCGATG TGGAGAAAGA ACTCAAAAAG ATGTCCCAAC GCGGAGAGAA AGATGAACTT
                                                                            840
      CGSTTCGTAT CGCTTTCTCA GGTTCTGGCC AATGGCCCGA TGAACAAAAC GAAAGGCAGT
                                                                            900
      CGGATCGCCG TTCTCTTTGC CGAAGGTGAA ATAACGGAAG AAATAATAAA GAAGCCGTTC
                                                                            960
      GACACTGACG GTAGCTCCAT CACACAAGAA CTCGCCAAAG AAATCAAGGC AGCAGCCGAT
                                                                           1020
10
      GACGATGATA TCAAAGCCGT AGTACTTCGT GTCAATTCTC CGGGAGGTAG TGCTTTCACT
      TCCGAACAGA TATGGAAGCA GGTAGCCGAT CTCAAGGCCA AAAAGCCTAT CGTGGTCTCC
                                                                           1140
      ATGGGCGACG TAGCAGCCTC GGGCGGATAC TACATAGCCT GCGCAGCCAA CAGTATCGTG
      GCAGAGCATA CGACTCTGAC CGGCTCCATC GGCATATTCG GCATGTTCCC GAACTTCGCG
                                                                           1260
      GGCGTAGCCA AGAAGATAGG AGTGAATATG GACGTCGTAC AGACATCCAA GTATGCAGAC
                                                                           1320
15
      TTGGGCAACA CCTTCGCTCC GATGACGGTC GAAGATCGTG CCCTCATCCA ACGCTACATA
                                                                           1380
      GAGCAGGGCT ACGACCTCTT CCTCACTCGC GTATCGGAAG GCCGCAACCG CACCAAGGCA
                                                                           1440
      CAGATCGACA GCATCGCTCA AGGCCGTGTA TGGCTCGGCG ACAAAGCTCT TGCACTCGGT
                                                                           1500
      TTGGTGGATG AGCTTGGAGG TTTGGACACA GCTATCAAAC GGGCCGCGAA GCTGGCTCAG
                                                                           1560
      CTCGGTGGCA ACTACAGCAT AGAGTATGGC AAGACCAAGC GCAACTTCTT CGAAGAGTTG
                                                                           1620
20
      CTCTCCTCAT CAGCAGCGGA TATGAAGTCT GCCATCCTGA GTACCATTCT CTCCGATCCG
                                                                           1680
      GAAATAGAAG TTCTGCGCGA ACTCCGCTCC ATGCCGCCCC GTCCTTCGGG CATACAGGCA
                                                                           1740
      CGTCTCCCCT ATTACTTCAT GCCGTAC
                                                                           1767
25
      (2) INFORMATION FOR SEQ ID NO:136
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 939 base pairs
                (B) TYPE: nucleic acid
30
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
35
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
40
                (A) ORGANISH: PORYPHYROHONAS GINGIVALIS
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...939
45
          (xi) SEQUENCE DESCRIPTION: SEO ID NO:136
      ATGAGAGCAA ACATTTGGCA GATACTTTCC GTTTCGGTTC TCTTTTCTT CGGGACAGCG
                                                                             60
      ATCGGACAGG CTCAGAGTCG AAACCGTACA TACGAGGCTT ATGTGAAACA GTACGCCGAC
                                                                            120
50
      GAAGCTATCC GACAGATGAG CCGCTACAAT ATACCGGCAA GCATCACCAT AGCACAGGCT
                                                                            180
      TTGGTGGAGA CAGGAGCCGG AGCCAGTACA CTGGCCAGCG TACACAACAA TCACTTCGGG
                                                                            240
      ATCAAATGCC ACAAATCGTG GACGGGCAAG CGCACCTATC GTACCGACGA TGCGCCGAAC
                                                                            300
      GAATGCTTCC GCAGCTATTC GGCCGCTCGC GAATCGTATG AAGATCATTC CCGATTTCTG
                                                                            360
      420
55
      ACGGGGTTGC AACGCTGTGG CTATGCCACC AATCGGGGCT ATGCCAATCT GCTGATCAAG
                                                                            480
      ATGGTGGAGC TGTATGAGCT ATATGCTTTG GATCGCGAGA AGTACCCCTC ATGGTTCCAC
                                                                            540
      AAGTCTTACC CCGGGTCCAA CAAAAAATCC CATCAAACGA CCAAGCAGAA GCAGAGCGGA
                                                                            600
      CTCAAGCACG AAGCTTACTT CAGCTACGGA CTGCTCTACA TCATAGCCAA GCAAGGCGAT
                                                                            660
      ACCTTCGATT CTTTGGCCGA AGAGTTCGAC ATGAGAGCCT CCAAACTGGC CAAATACAAC
                                                                            720
60
      GATGCTCCCG TGGATTTCCC GATCGAAAAG GGCGATGTGA TCTATCTGGA GAAAAAGCAC
                                                                            780
      GCATGCTCCA TCTCCAAACA CACACAGCAC GTAGTGCGTG TGGGCGATTC GATGCACAGT
                                                                            840
      ATCTCCCAAC GCTATGGCAT CCGGATGAAG AACCTCTACA AGCTCAACGA CAAGGATGGC
                                                                            900
      GAATATATAC CCCAAGAGGG CGATATACTG CGCTTGCGC
                                                                            939
65
      (2) INFORMATION FOR SEQ ID NO:137
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1569 base pairs
70
                (B) TYPE: nucleic acid
                (C) STRANDEDHESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
75
```

```
(iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
 5
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
10
                   (B) LOCATION 1...1569
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:137
       ATGGACGGAC GTCGATATTC GGATGGCCTC CATCAGGCTA TCGAAGCCAA AGAGCATGTG
15
       AAAGTAGAGG CTGCGACACA GACATTTGCA ACTATCACTT TGCAGAACTA TTTCCGCATG
                                                                                      120
       TATCATAAGC TGGCAGGGAT GACCGGTACT GCTGAAACTG AAGCGGGAGA GCTTTGGGAC
                                                                                      180
       ATCTACAAAC TGGACGTTGT AGTTATTCCG ACAAACAAGC CTATCGCCCG TAAGGATATG
       AATGATCGTA TCTATAAGAC GGCACGTGAA AAATATGCAG CAGTTATCGA AGAGATTGTA
                                                                                      300
       CGTCTTGTCG AAGAGGGCAG ACCTGTACTT GTCGGTACTA CTTCGGTGGA AATATCCGAA
                                                                                      360
20
       TTGTTGAGCC GTATGTTACG CTTGCGTGGC ATCCAACACA ATGTACTCAA TGCCAAATTG
                                                                                      420
       CATCAGAAGG AGGCCGAGAT TGTAGCTCAG GCCGGTCAGA AAGGAACTGT TACCATCGCA
                                                                                      480
       ACGAACATGG CCGGTCGTGG TACCGACATC AAGCTCTCTG CCGAGGTTAA GAAAGCCGGG
GGTTTGGCTA TCATTGGTAC GGAAAGGCAC GAATCCAGAC GAGTGGACAG ACAGCTTCGT
                                                                                      540
                                                                                      600
       GGTCGTTCCG GCCGTCAGGG TGATCCCGGT TCGTCCATAT TCTATGTTTC CCTTGAAGAT CATCTGATGC GCCTCTTTGC CACAGAAAAG ATTGCATCAT TGATGGATCG TTTAGGTTTC
                                                                                      660
25
                                                                                      720
       AAGGAAGGAG AAGTGCTCGA AAACAACATG CTGAGTAAGT CCGTGGAGCG TGCTCAAAAG AAGGTGGAAG AGAACAACTT CGGTATCCGT AAACATCTGC TTGAGTACGA TGATGTAATG
                                                                                      780
                                                                                      840
       AATTCGCAGC GTGAAGTCAT TTATACCCGT CGCCGTCATG CTTTGATGGG AGAGCGTATC
                                                                                       900
       GGTATGGATG TACTCAATAC CATATACGAC GTATGTAAGG CTCTGATTGA CAATTATGCA
                                                                                      960
30
       GAAGCCAATG ATTTCGAAGG CTTCAAGGAA GATCTGATGC GTGCACTCGC GATAGAATCT
                                                                                     1020
       CCTATCACGC AAGAAATATT CAGAGGTAAG AAAGCAGAAG AGCTGACCGA TATGCTTTTC
                                                                                     1080
       GATGAAGCTT ACAAGTCTTT CCAACGTAAG ATGGATCTGA TCGCAGAAGT GGCCCACCCT
                                                                                     1140
       GTGGTTCATC AGGTATTCGA GACCCAAGCC GCCGTGTACG AGCGCATTCT AATCCCCATT
                                                                                     1200
       ACGGATGGTA AACGTGTCTA TAACATAGGA TGCAATTTGC GTGAAGCGGA TGAAACTCAA
                                                                                     1260
35
       GGGAAAAGCA TCATCAAAGA ATTTGAGAAA GCTATCGTAC TGCATACTAT CGATGAGTCT
                                                                                     1320
       TGGAAAGAAC ATCTGCGTGA GATGGACGAG CTTCGTAATT CCGTTCAGAA TGCCAGCTAC
                                                                                     1380
       GARAACAAAG ATCCACTACT TATCTATAAA CTCGAATCTT ACGAACTGTT CCGCAAGATG
GTAGAAGCCA TGAACCGTAA GACCGTAGCG ATCCTAATGC GTGCTCGGAT ACCGGTACCG
                                                                                     1440
                                                                                     1500
       GAGGCTCCTT CCCAAGAAGA GCTGGAACAC AGGCGGCAAA TAGAAATCCG ACATGCAACC
                                                                                     1560
40
       CAACAACGT
                                                                                     1569
       (2) INFORMATION FOR SEQ ID NO:138
45
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1125 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDHESS: double
                   (D) TOPOLOGY: circular
50
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
55
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
60
           (ix) FEATURE:
                  (A) MAME/KEY: misc feature
                   (B) LOCATION 1...1125
           (xi) SEQUENCE DESCRIPTION: SEO ID NO:138
65
       ATGAATTTCT TAAAAAAAGA ACCGTTTAAA ATATTCTCTA TGATTTATCT GCTGTTAGAT
       ACAATAACAA ACCGTGCCGG TACAGAACGC GCCGTGATCA ACTTGGCTAA CAACCTGCAT GCCAATGGTC ATCGCGTATC ATTAGTCAGC GTTTGTACAA AAGAAGGAGA GCCTTCCTTC
                                                                                      120
                                                                                      180
       CAAGTAGAAA AAGGAATAGA AGTACACCAT CTCGGAATTA GGCTTTATGG CAATGCATTA
                                                                                      240
70
       GCCCGCAAAA CAGTATATTT CAAGGCTTAT CGAAGGATAA AAGCCCTATA CAAGAAGCGT
                                                                                      300
       GAACCGGTTT TATTGATAGG GACTAATATT TTTATCAATA CAATTTTGTC TCAGATCAGT
                                                                                      360
       AACAGAGGCA GAATATTTAC GATCGGATGC GAACATATCT CTTATGATAT TGCCCGCCCT
                                                                                      420
       ATTACAAAAC GCATAAGGGG GTTTCTGTAT TCAGGGCTTG ATGCCGTTGT AGCACTGACA
                                                                                      480
       AAAAGAGATC AGCAATCGTT CGAGGCAATC TTACGTGGAC GCTCTAAAGC ATATGTCATA
                                                                                      540
75
       CCCAATCAAG TTTCATTTAC TACAGTCCAA AGAGATGCTA CTACTCACAA ACAAATGTTG
                                                                                      600
```

```
GCGATTGGCA GGCTTACCTA CCAGAAGGGT TTTGAATTCA TGATAGAAGA TGCATCACGA
                                                                                660
      GTGCTGCGAG AAAGGCCTGA TTGGAAGCTT ATCATAGTCG GAGATGGCGA AAATGAATCG
                                                                                720
      ATGCTACGTA AAGAAATTGC ATCTCGCAAT ATGGAGTCGC AAATAGAAAT ACATCCATCT
                                                                                780
      ACACCGGAAA TTCGCAAATA CTACGAATCA TCTGCTATTT ATCTAATGAC GTCCCGTTTC
                                                                                840
 5
      GAAGGACTAC CAATGGTACT TCTCGAAGCA GAAGCATATG CACTACCTAT AATCTCATAC
                                                                                900
      GATTGTCCGA CCGGCCCGAG GGAACTGATC GAAAACGGTC GCAATGGTTT CCTTGTGCCA
                                                                                960
      ATGGAAGCAC ATGAAGACTT CGCGGATAAG TTACGCTTAT TGATGGATGA TGAAACTCTT
                                                                               1020
      CGTAAGAAAA TGGGACAAGA ATCAGAGTTG ATGGTCAAAT CCTACTCTCC GGCAAATATC
                                                                               1080
      TATGAATGTT GGAAGAAACT ATTCGTCGAA ATCGGCTACA TGAAT
                                                                               1125
10
      (2) INFORMATION FOR SEQ ID NO:139
            (i) SEQUENCE CHARACTERISTICS:
15
                 (A) LENGTH: 1086 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
20
           (ii) HOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
25
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
30
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1086
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:139
35
      ATGATTTATC TGCTGTTAGA TACAATAACA AACCGTGCCG GTACAGAACG CGCCGTGATC
      AACTTGGCTA ACAACCTGCA TGCCAATGGT CATCGCGTAT CATTAGTCAG CGTTTGTACA
                                                                                120
      AAAGAAGGAG AGCCTTCCTT CCAAGTAGAA AAAGGAATAG AAGTACACCA TCTCGGAATT
                                                                                180
      AGGCTTTATG GCAATGCATT AGCCCGCAAA ACAGTATATT TCAAGGCTTA TCGAAGGATA
                                                                                240
      AAAGCCCTAT ACAAGAAGCG TGAACCGGTT TTATTGATAG GGACTAATAT TTTTATCAAT
                                                                                300
40
      ACAATTTTGT CTCAGATCAG TAACAGAGGC AGAATATTTA CGATCGGATG CGAACATATC
                                                                                360
      TCTTATGATA TTGCCCGCCC TATTACAAAA CGCATAAGGG GGTTTCTGTA TTCAGGGCTT
                                                                                420
      GATGCCGTTG TAGCACTGAC AAAAAGAGAT CAGCAATCGT TCGAGGCAAT CTTACGTGGA
                                                                                480
      CGCTCTAAAG CATATGTCAT ACCCAATCAA GTTTCATTTA CTACAGTCCA AAGAGATGCT
                                                                                540
      ACTACTCACA AACAATGTT GGCGATTGGC AGGCTTACCT ACCAGAAGGG TTTTGAATTC
                                                                                600
45
      ATGATAGAAG ATGCATCACG AGTGCTGCGA GAAAGGCCTG ATTGGAAGCT TATCATAGTC
                                                                                660
      GGAGATGGCG AAAATGAATC GATGCTACGT AAAGAAATTG CATCTCGCAA TATGGAGTCG
                                                                                720
      CAAATAGAAA TACATCCATC TACACCGGAA ATTCGCAAAT ACTACGAATC ATCTGCTATT
                                                                                780
      TATCTAATGA CGTCCCGTTT CGAAGGACTA CCAATGGTAC TTCTCGAAGC AGAAGCATAT
                                                                                840
      GCACTACCTA TAATCTCATA CGATTGTCCG ACCGGCCCGA GGGAACTGAT CGAAAACGGT
                                                                                900
50
      CGCAATGGTT TCCTTGTGCC AATGGAAGCA CATGAAGACT TCGCGGATAA GTTACGCTTA
TTGATGGATG ATGAAACTCT TCGTAAGAAA ATGGGACAAG AATCAGAGTT GATGGTCAAA
                                                                                960
                                                                               1020
      TCCTACTCTC CGGCAAATAT CTATGAATGT TGGAAGAAAC TATTCGTCGA AATCGGCTAC
                                                                               1080
      ATGAAT
                                                                               1086
55
      (2) INFORMATION FOR SEQ ID NO:140
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1920 base pairs
60
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
65
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
70
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KET: misc feature
75
                 (B) LOCATION 1...1920
```

```
ATGGGAAAAA TCATTGGAAT TGACTTAGGC ACAACGAACT CTTGTGTCTC TGTATTGGAA
 5
      GGTAACGAAC CTATCGTTAT TACAAACAGT GAGGGCAAGC GCACAACGCC CTCGGTAGTG
                                                                               120
      GCTTTTGTGG ATGGTGGCGA GCGTAAGGTG GGCGATCCGG CCAAGCGTCA GGCCATCACC
                                                                               180
      AATCCGACCA AGACGATATA CTCTATCAAA CGCTTCATGG GCGAAACTTA CGATCAGGTT
                                                                               240
      TCCAGAGAAG TGGAGAGAGT GCCATTCAAG GTAGTACGTG GGGACAATAA TACTCCGCGC
                                                                               300
      GTAGATATAG ACGGTCGTCT CTATACGCCG CAGGAAATTT CGGCCATGAT CCTTCAGAAG
                                                                               360
10
      ATGAAGAAGA CGGCCGAAGA CTACCTCGGT CAGGAAGTAA CGGAGGCCGT GATCACTGTG
                                                                               420
      CCCGCATACT TCAACGACGC TCAACGTCAG GCAACGAAAG AAGCAGGAGA GATCGCCGGC
                                                                               480
      CTGAAAGTTC GCCGTATTGT GAACGAGCCT ACGGCAGCTT CTCTGGCCTA CGGTCTGGAC
                                                                               540
      AAGTCCAATA AGGATATGAA GATCGCTGTC TTCGACTTGG GTGGCGGTAC CTTCGATATC
                                                                               600
      TCTATCTTGG AATTGGGCGA CGGCGTTTTC GAAGTGAAAT CGACCAACGG TGATACGCAC
                                                                               660
15
      CTCGGAGGAG ACGACTTCGA CCACGTGATC ATTGACTGGC TGGCAGAAGA GTTCAAGTCT
                                                                               720
      CAGGAAGGTG TGGATCTTCG CCAGGATCCT ATGGCTATGC AGCGTCTGAA AGAAGCTGCC
                                                                               780
      GAAAAAGCCA AGATAGAGCT CTCCAGCACT TCATCTACGG AGATCAACCT CCCCTATATC
      ATGCCGGTGA ACGCCATCCC CAAGCACTTG GTGATGACGC TTACAAGGGC TAAGTTCGAG
      CAGTTGGCCG ATCGTCTGAT TCAGGCATGT GTGGCACCCT GCGAAACGGC CTTGAAAGAT
                                                                               960
20
      GCCGGTATGT CACGTGGCGA TATCGATGAA GTGATTCTCG TAGGTGGTTC CACACGTATT
                                                                              1020
      CCTGCTATTC AGGAGATTGT GGAGAAGATC TTCGGTAAGG CTCCGTCCAA GGGTGTGAAT
                                                                              1080
      CCCGACGAAG TGGTAGCTGT GGGTGCCGCT ATTCAAGGCG GTGTTCTGAC CGGTGAGGTA
AAGGATGTCT TGCTGTTGGA CGTTACCCCC TTGTCGCTCG GTATCGAGAC TATGGGAGGC
                                                                              1140
                                                                              1200
      GTGATGACTC GCTTGATCGA TGCCAATACC ACTATCCCGA CGAAGAAGAG CGAAATCTTT
                                                                              1260
25
      ACCACAGCAG TGGACAATCA ACCTTCGGTA GAGATTCATG TACTTCAGGG TGAGCGTTCT
                                                                              1320
      TTGGCTAAGG ACAATAAGAG CATCGGCCGT TTCAACTTGG ACGGTATTGC TCCGGCGCCC
                                                                              1380
      CSTCAGACAC CGCAGATCGA AGTAACGTTT GACATCGATG CCAACGGTAT CCTGAATGTA
                                                                              1440
      ACGGCTCATG ACAAAGCTAC CGGCAAGAAG CAGAATATCC GCATCGAAGC CTCCAGCGGT
                                                                              1500
      TTGTCCGATG ATGAGATCAA GCGCATGAAG GAAGAGGCGC AGGCCAATGC CGAAGCAGAT
                                                                              1560
30
      AAGAAAGAGA AAGAACGTAT CGACAAGATC AATCAGGCCG ACAGCATGAT CTTCCAGACG
      GAAAAGCAGT TGAAGGAGTT GGGAGACAAA TTCCCGGCCG ACAAGAAGGC TCCGATCGAT
                                                                              1680
      ACCGCTCTCG ACAAACTGAA AGAAGCACAC AAAGCACAGG ATGTAGCTGC TATCGATACA
                                                                              1740
      GCCATGGCCG AACTGCAAAC CGCTCTTTCC GCAGCGGGCG AAGAGCTTTA CAAGAATGCC
                                                                              1800
      GGAGCAGCCC AAGGTGGCGC ACAACCCGGT CCGGACTTCG GCGGTGCTCA AGGTCCCTCT
                                                                              1860
35
      GCCGGTGATC AGCCCTCTGA CGACAGAAC GTCACAGACG TAGACTTCGA GGAAGTGAAG
                                                                              1920
      (2) INFORMATION FOR SEQ ID NO:141
40
            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1347 base pairs
                 (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
45
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
50
          (iv) ANTI-SEUSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROHONAS GINGIVALIS
55
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1347
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141
60
      ATGCGCTACG ACTTAGCTAT CATCGGTGGA GGGCCGGCCG GTTATACGGC TGCCGAACGT
                                                                                60
      GCTGCCAAAG GTGGCCTGAA AACCCTCCTA ATTGAGAAGA ATGCTCTCGG TGGTGTATGC
                                                                               120
      180
                                                                               240
65
      GTGATTGCCA GAAAAGGTAA AATCATTCGC AAGCTGACTG CAGGCATCCG TTCACGCCTG
                                                                               300
      ACAGAGGCCG GAGTAGAGAT GGTGACGGCA GAAGCTACCG TAACGGGATG CGATGCAGAC
                                                                               360
      GGCATCATCG GCATTACTGC GGGCGAAGCA CAGTACAAAG CTGCCAACCT GCTACTATGT
                                                                               420
      ACCGGTTCGG AGACGTTTAT TCCACCCATC CCCGGAGTGG AGCAGACAGA GTATTGGACA
                                                                               480
      AACCGTGAAG CTCTACAGAA CAAAGAGATT CCGACCTCTC TCGTCATCAT CGGTGGTGGA
70
      GTGATCGGAA TGGAGTTCGC TTCTTTCTTC AACGGTATCG GTACGCAAGT GCACGTGGTG
      GAGATGCTGC CGGAAATACT CAACGGTATC GATCCCGAAC ATGCAGCTAT GCTACGCGCT
                                                                               660
      CACTATGAAA AAGAAGGAAT CAAATTCTAC CTCGGGCACA AAGTAACATC GGTTCGCAAC
                                                                               720
      GGAGCTGTTA CGGTAGAATA CGAAGGAGAA AGCAAAGAGA TCGAAGGAGA ACGTATCCTG
                                                                               780
      ATGAGTGTGG GACGTCGCCC CGTGCTGCAA GGATTCGAGT CGCTCGGATT GGTGCTTGCC
GGCAAAGGTG TAAAGACTAA TGAGAGGATG CAAACTTCCC TGCCCAATGT CTATGCTGCA
                                                                               840
75
                                                                               900
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140

```
GGTGATATTA CAGGCTTCTC GCTTTTGGCA CATACGGCTG TACGGGAAGC AGAGGTAGCA
      GTAGATCAGA TTTTGGGCAA AACAGACGAA ACGATGAGCT ACCGTGCCGT ACCAGGTGTG
      GTGTACACCA ATCCCGAGGT CGCCGGTGTG GGAGAGACGG AAGAATCGCT TCGCAAAGCA
      GGACGTGCCT ACACTGTTCG TCGCCTTCCT ATGGCCTTCT CCGGTCGATT TGTAGCAGAA
                                                                            1140
 5
      AACGAACAAG GCAATGGAGA GTGCAAACTA CTACTTGATG AAGAGAACCG CTTGATCGGA
                                                                            1200
      GCACACCTCA TTGGCAATCC GGCCGGCGAA CTCATCGTAA CCGCTGCCAT GGCCATCGAG
                                                                            1260
      ACCGGCATGA CGGATCGACA AATCGAACGA ATCATATTCC CTCATCCGAC TGTAGGCGAA
                                                                            1320
      ATCCTAAAAG AAACTCTCGC CGGAGGT
                                                                            1347
10
      (2) INFORMATION FOR SEQ ID NO:142
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 2823 base pairs
15
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DMA (genomic)
20
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
25
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORTPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
30
                (B) LOCATION 1...2823
          (xi) SEQUENCE DESCRIPTION: SEO ID NO:142
      ATGGAATTGA AAAGATTTTT ATCACTTGGT CTTCTGCTTG TGGGATTCAT TCCGATGAAG
                                                                              60
35
      CTTTCTGCCC AACAGGCTCA GCCACTCCCT ACAGATCCGG CTGTTCGTGT CGGTAAGTTG
                                                                             120
      GACAACGGAT TGACTTATTT CATCCGTCAC AACGAGAACC CGAAAGATCG TGCGGATTTC
                                                                             180
      TTTATCGCAC AAAAGGTAGG TTCTATTCTT GAAGAAGATA GCCAGTCCGG TTTGGCTCAC
                                                                             240
      TTCTTGGAAC ACATGGCTTT CAACGGTACG AAGAACTTCC CCGGTAAGAA CTTGATCAAC
                                                                             300
      TATCTCGAAA CGATCGGTGT ACGTTTCGGT CAGAACCTGA ACGCTTCTAC CGGATTCGAC
                                                                             360
40
      AAGACGGAAT ATACGATAAT GGATGTGCCG ACTACACGTC AGGGAATCAT CGACTCCTGC
      TTGCTTATCC TGCATGATTG GAGTAACAAT ATTACCCTCG ACGGGCATGA GATCGACGAG
                                                                             480
      GAGCGCGGTG TGATCCAGGA AGAGTGGCGT GCTCGTCGCG ATGCCAACCT TCGTATGTTC
      GAGGCTATAC TTGCCAAGGC TATGCCGGGT AATAAATATG CAGAACGCAT GCCCATCGGT
                                                                             600
      CTGATGGACG TCGTGCTCAA CTTCAAGCAT GATGAGCTGC GCAACTATTA TAAGAAATGG
45
      TATCGTCCCG ACCTGCAAGG TCTGGTGATC GTGGGAGATA TCGATGTGGA CTATGTGGAG
      AACAAGATCA AAGAACTCTT CAAGGACGTT CCTGCTCCCG TGAATCCAGC AGAGCGTATC
      TATACGCCGG TAGAGGACAA CGATGAGCCT ATCGTAGCCA TTGCTACCGA TGCTGAGGCT
                                                                             840
      ACTACCACGO AGCTCTCCAT CAGCTTCAAG AGCGACCCCA CTCCTCAAGA AGTGCGAGGA
      TCGATATTCG GACTTGTGGA AGACTATATG AAACAGGTGA TCACTACAGC CGTGAATGAG
                                                                             960
50
      CGTCTGTCCG AGATTACTCA CAAGCCTAAC GCTCCTTTCC TCAGTGCAGG AGCTTTCTTC
      TCTAACTTCA TGTACATCAC CCAGACTAAG GACGCATTCA ATTTTGTTGC CACGGTTCGT
                                                                            1080
      GAGGGTGAAG CGGAGAAAGC GATGAACGCA TTGGTGGCAG AGATAGAAAG CCTCCGTCAG
                                                                            1140
      TTCGGTATCA CCAAAGGCGA ATACGATCGT GCACGCACGA ATGTGCTCAA GCGATACGAG
                                                                            1200
      AATCAATACA ACGAAAGAGA CAAGCGTAAG AACAATGCTT ATGCCAATGA ATACTCCACC
                                                                            1260
55
      TACTTCACCG ATGGCGGCTA TATCCCGGGT ATTGAGGTGG AATATCAGAC GGTGAATGCT
                                                                            1320
      TTTGCTCCTC AGGTTCCTCT GGAAGCATTC AATCAGGCTA TTGCCCAAAT GATCGATCCG
                                                                            1380
      GTGAAGAATG CTGTCGTTAC CCTCACCGGT CCTTCAAAGG CTGAAGCCAA GATTCCGAGC
                                                                            1440
      GAAGCAGACT TCCTCGCTGC TTTCAAAGCT GCTCGTCAGC AGAAAGTAGA AGCCAAGAAA
                                                                            1500
      GACGAAGTCT CCGACCAAAA ATTGATGGAG AAAGCTCCTA AGGCCGGAAA GATCGTTTCC
                                                                            1560
60
      GAGAAGAAG ATCAGAAGTT CGGTACCACG GAACTTACCC TTAGCAATGG CATCAAAGTA
                                                                            1620
      TACCTCAAGA AGACCGATTT CAAATCAAAC GAAATCCTGA TGAGTGCTCT CAGCCCGGGT
                                                                            1680
      GGTATCCTCT CCGGAAAGCA TGCTCCCAAC CAATCTGTGA TGAATTCGTT CATGAACGTG
                                                                            1740
      GGTGGCTTGG GCAACTTCGA TGCTATCCAG CTGGATAAGG TGCTGACAGG TCGCTCTGCT
                                                                            1800
      TCCGTATCTC CCTCTTTGTC TCTGCTCAGT GAAGGTCTTT CGGGCAAAAC GACTGTAGAA
                                                                            1860
65
      GATATGGAAA CTTTCTTCCA GTTGATCTAT CTCCAAATGA CTGCTAACCG CAAGGATCCC
                                                                            1920
      GAAGCGTTCA AGGCCACACA GGAAAAGTTG TACAATAACT TGAAAAATCA GGAAGCCAAC
                                                                            1980
      CCGATGGCTG CGCTTATGGA CTCTATCCGT CATACCATGT ACGGCGATAA TCCGATGATG
                                                                            2040
      AAACCCATGA AAGCTGCTGA CGTGGAGAAA GTAAATTACG ATCAGGTAAT GGCTTTCTAC
                                                                            2100
      AATGAGCGAT TCGCTGATGC CGGCGACTTT ATGTTCTTCT TTATCGGTAA TCTGGATGAA
                                                                            2160
70
      GCCAAGATGA AGCCATTGAT CGAAACTTAT CTTGCTTCAT TGCCCAACCT CAAGCGTGGC
                                                                            2220
      GATAAGATGA ATAAGGCTCA GGTACCGGCT GCCCGTTCGG GAAAGATCGA TTGCAAGTTC
                                                                            2280
      GAGAAGGAAA TGGATACTCC TTCGACTACT ATATTCGATG TCGTGTCCGG AAATGTGGAA
                                                                            2340
      TATACGCTCA AGAACAGTCT CCTGCTGGAA GTCTTCTCAG CCGTAATGGA TCAGGTGTAC
                                                                            2400
      ACGGCTACCG TTCGCGAGAA GGAAGGCGGT GCATACAGTG TGGCTGCATT CGGCGGTCTC
                                                                            2460
75
      GAGCAATATC CTCAGCCCAA GGCTCTGATG CAGATCTATT TCCCCACGGA TCCTGCTCGT
                                                                            2520
```

```
GCCGAGGAAA TGAATGCTAT CGTTTTTGCT GAGTTGGAGA AGCTTGCCAA GGAGGGCCCC
                                                                              2580
      AATGTGGAAT ACTTTAAGAA GACTATCGAA AACCTGAATA AGCAGCACAA AGAAAGTCTG
                                                                              2640
      CGTGAGAATC GTTTCTGGCT CGAAGCCATG AAGGCGTCTT TCTTCGAAGG AAATGACTTC
                                                                              2700
                                                                              2760
      ATCACAGACT ACGAATCCGT ACTGAACGGT CTTACTCCTG CTGAATTGCA AAAGTTTGCG
 5
      GCAGACCTCT TGAAGCAGCA GAATCGGGTT GTTGTCATGA TGGCTCCTGT TGCAAAGGCT
                                                                              2820
                                                                              2823
      (2) INFORMATION FOR SEQ ID NO:143
10
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 2052 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
15
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DHA (genomic)
         (iii) HYPOTHETICAL: NO
20
          (iv) AUTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMOMAS GINGIVALIS
25
          (ix) FEATURE:
                (A) HAME/KEY: misc_feature
                (B) LOCATION 1...2052
30
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:143
      ATGAGTAAGA AAGGAACAAT CGGGGTAACG AGCGACAATA TATTCCCCGT CATCAAAAAA
      TTCCTGTACA GCGACCATGA GATATTCCTG CGTGAGATCG TCTCCAATGC CGTGGATGCT
                                                                               120
      ACGCAGAAGC TGAAAACGCT TACATCCGTC GGCGAATTCA AAGGCGAGAC GGGTGACCTC
                                                                               180
35
      CGCGTAACGG TCAGCGTGGA TGAAGTGGCA CGCACGATCA CGGTCAGCGA CCGCGGCGTA
                                                                               240
      GGGATGACCG AAGAGGAGGT GGAGAAGTAC ATCAATCAGA TTGCTTTCTC CAGTGCGGAA
                                                                               300
      GAGTTTCTTG AAAAGTACAA AGACGACAAG GCCGCCATTA TCGGCCACTT CGGACTCGGA
                                                                               360
      TTTTACTCGG CTTTCATGGT GTCCGAGCGA GTGGACGTGA TCACGCGCTC TTTCCGAGAA
                                                                               420
      GATGCTACGG CGGTGAAATG GAGCTGCGAC GGATCGCCCG AATACACGCT CGAACCTGCG
                                                                               480
40
      GACAAGGCTG ACCGTGGCAC CGACATCGTG ATGCACATCG ATGAGGAGAA TAGCGAGTTC
                                                                               540
      CTCAAAAAAG AAAAGATAGA GGGGCTCCTC GGCAAATACT GTAAGTTCCT TACCGTGCCG
                                                                               600
      ATCATTTCG GCAAGAAGCA GGAATGGAAA GACGGCAAGA TGCAAGATAC GGACGAGGAC
                                                                               660
      AATCAGATCA ACGACACACA TCCTGCCTGG ACCAAAAAGC CTGCCGACCT CAAGGACGAA
                                                                               720
      GACTATAAGG AATTTTACCG TTCGCTCTAT CCCATGTCCG AAGAGCCTCT CTTCTGGATC
                                                                               780
45
      CACCTCAATG TGGACTATCC GTTCAATCTG ACAGGTATCC TCTATTTCCC GAAGATCAAA
                                                                               840
      AACAACTTGG ATCTGCAGCG CAACAAGATT CAGCTCTACT GCAATCAGGT TTACGTCACC
                                                                               900
      GATGAAGTAC AGGGTATCGT GCCGGACTTC CTCACCCTCC TGCACGGGGT CATCGATTCG
                                                                               960
      CCGGATATTC CCCTCAACGT ATCGCGCTCC TATCTGCAGA GCGATGCCAA TGTGAAGAAG
                                                                              1020
      ATCTCGTCTC ATATCACCAA GAAGGTGGCA GACCGTCTGG AAGAAATTTT CAAAAACGAC
                                                                              1080
50
      CGCCCCACAT TCGAGGAGAA ATGGGATAGT CTGAAGCTCT TCGTCGAATA CGGTATGCTG ACGGATGAGA AGTTCTATGA GCGTGCAGCC AAATTCTTCC TTTTCACCGA TATGGACGGA
                                                                              1140
                                                                              1200
      CACAAGTACA CGTTCGACGA ATACCGAACG CTCGTCGAAG GTGTACAGAC GGATAAGGAC
                                                                              1260
      GGACAGGTAG TGTATCTCTA TGCTACGGAC AAGCATGGAC AGTACAGCCA CGTGAAACGT
                                                                              1320
      GCATCCGACA AAGGCTACAG CGTGATGCTG TTGGATGGTC AGTTGGATCC GCATATCGTG
                                                                              1380
55
      AGCCTGCTGG AGCAAAAGTT GGAGAAGACA CACTTTGTCC GTGTCGATAG CGATACGATC
                                                                              1440
      AACAATCTGA TCCGCAAGGA GGAAAGAGCC GAAGTGAAAC TGTCCGATAC GGAGCGCGCC
                                                                              1500
      ACTCTCGTGA AGCTGTTCGA AGCACGCCTG CCACGGGACG AGAAGAAGCA CTTCAATGTA
                                                                              1560
      GCTTTCGAAT CGCTCGGAGC CGAAGGTGAA GCCATCCTTA TCACACAAGC CGAATTCATG
                                                                              1629
      CGCCGTATGC GCGATATGGC ACAGCTGCAG CCGGGAATGA GCTTCTACGG CGAACTCCCC
                                                                              1680
60
      GATTCGTACA ATCTGGTACT TAATACCGAT CATCCGCTCA TCGACAGGGT ACTCTCCGGT
      GAGAAAGAAT CGGTAGAGCC TTCGCTCACA GAGCTTAGAG CGAAAATCGC CGAGCTGAAA
                                                                              1800
      GCGGAAGAGG CCAAGCTGCT CGATGAGGAA AAAGGGAAGA AACCGGAGGA AATCCCTGTT
                                                                              1860
      GCCACGAAGG AAGCCAAGGA GAACAACGCC GTCGAACAGG CCAAAACCGA AGGCAGTATC
                                                                              1920
      AACGATCAAC TGACCAAATA TGCTCAGGAC AACGAGCTGA TAGGTCAGCT CATCGACTTG
                                                                              1980
65
      GCTCTGCTCG GAAGCGGATT GCTGACGGGA GAGGCTTTGG CCGAATTCAT TCGTCGCAGC
                                                                              2040
      CAGCGTCTTC TC
                                                                              2052
      (2) INFORMATION FOR SEQ ID NO:144
70
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1392 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
75
                (D) TOPOLOGY: circular
```

```
(ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
 5
           (iv) AUTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
10
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1392
15
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:144
      ATGGAAAAAC TGATCGATAT TTTGGTCGTA GACGATGATG TGGCAGTCTG TGCCGCACTG
                                                                                60
      CGTCTGGTGC TCAAGCGAGC GGGCTATAAT CCCGTTATAG CCAACAGTCC CGACGAAGCT
                                                                               120
      TTGTCCATAA TGCGGAATCC TGATGGCGGC TGTAAGCCGG CTGTGATTCT GATGGATATG
                                                                               180
20
      AATTTCTCCC TTTCGACCTC CGGCAGGGAA GGATTGGAAC TACTGGAGAA GATGCAGATA
                                                                               240
      TTCACTTCCT GCCCTGTCAT ACTGATGACG GCTTGGGCTT CGATTCCACT GGCAGTGGAG
                                                                               300
      GGAATGAGGC TTGGAGCTTT CGACTTCATA GGCAAGCCAT GGGACAACGA TCGGCTCCTT
                                                                               360
      CGTACCATAG ATACGGCCTT GCATCTGGCT GCTCCCTCAG CTGTGGCGAA TCCATCGGAA
                                                                               420
      CAGTCTGACA GAGATACAGC CCGTCAGCCG AAAGCTACAG TCCAAGAGAA TGACCCCTGT
                                                                               480
25
      GCCCATATCA TAGGCCGGAG CGATGCCATC TGTAAGATCA AGGAACGGAT ACGCCGCATA
                                                                               540
      GCTCCCACCC ATGCCTCTGT GCTGATCACG GGCGAGAGCG GTACGGGCAA AGAGTTGATA
                                                                               600
      GCCGAAGCTC TGCACCGTGG GAGCAAACGA GCCTCAGCCC CATTCGTCAA GGTCAATTTG
                                                                               660
      GGTGGGATTC CCGAAAGTTT GTTCGAAAGT GAGCTGTTCG GACATAAGAA AGGAGCTTTT
                                                                               720
      ACCAATGCTT TTTCCGACAG GAAAGGACGG TTCGAGCTGG CTGATGGCGG CACGATCTTT
                                                                               780
30
      CTGGACGAAA TAGGCGAACT ACCGGTCGGC AACCAAGTAA AACTGCTGCG AGTGCTACAG
                                                                               840
      GAACAGACAT TCGAGCCGTT GGGCGAGAGC GTCTCCCACC GAGTGGACAT CCGTGTGGTA
                                                                               900
      TCGGCTACGA ATGCTTCCTT GGAGCGAATG GTAGCCGAAG GACGTTTCAG AGAGGACCTC
                                                                               960
      TACTATCGAA TCAACCTGAT ACATCTGCAT CTGCCTCCGC TGCGTGAGCG TCAGGAGGAT
                                                                              1020
      ATACAGCTGC TGGTGGAAGC CTTCAGTGAA GCCTTTGCCC AATCGAACGG ATTGCCCCAT
                                                                              1080
      GCCGTTTGGA GTGCGGAAGC TATGCGACGT ATCTGTGCCA TGCCCCTACC GGGCAATGTA
35
                                                                              1140
      CGCGAACTGA AAAACGTAGT GGAGCGTACG CTATTGCTCT CGGGATCGAG AGAAATCAGT GCCCGGGATG TGGCTGACTT CGGTTCGCAG GTGACGGCAG CAGACCACTC CGACGAACGG
                                                                              1200
                                                                              1260
      GCTTTGACCG ACATGGAGGA AGCTGCTATC CGAGAGACGC TGACTAAATA CAACGGCAAC
                                                                              1320
      GTTAGTCGTG CTGCACGAGC CTTGGGATTG AGCCGGGCAG CTCTTTACCG GCGAATGGAG
                                                                              1380
40
      AAATACGGAC TG
                                                                              1392
      (2) INFORMATION FOR SEQ ID NO:145
45
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 750 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
50
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
55
          (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
60
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...750
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:145
65
      ATGCTTAAGA TAAAGAACCT CCACGCCACA GTACAGGGCA AAGAGATATT GAAAGGAATC
      AATCTGGAGA TCAATGCCGG AGAGATTCAT GCTATCATGG GGCCGAACGG ATCGGGGAAA
      AGTACGCTCT CTTCCGTTTT GGTGGGACAT CCCTCCTTTG AAGTCACGGA AGGAGAGGTG
                                                                               180
      ACATTCAATG GAATCGACCT GCTCGAACTC GAACCGGAAG AACGTGCACA CCTCGGACTC
                                                                               240
70
      TITCTCAGTT TCCAATATCC GGTCGAGATC CCGGGCGTCA GCATGGTGAA TTTCATGAGG
                                                                               300
      GCAGCTGTCA ATGAACATAG GAAAGCGATC GGAGCAGAAC CCGTATCGGC AAGCGACTTC
                                                                               360
      CTCAAGATGA TGCGAGAGAA GCGTGCCATT GTGGAGCTGG ACAACAAATT GGCCAGCCGT
                                                                               420
      TCTGTGAACG AAGGCTTCTC CGGTGGAGAA AAAAAGAGGA ACGAAATCTT CCAAATGGCT
                                                                               480
      ATGCTCGAAC CCAAGCTGGC TATTTTGGAC GAAACCGATA GCGGGCTCGA TATCGACGCT
                                                                               540
75
      CTCCGCATCG TAGCAGGCGG GGTAAACCGA CTCCGCTCTC CGGAGAATGC TGCTATTGTG
                                                                               600
```

| | ATCACACACT ATCAGCGTTT GCTCGAGTAC ATCAAGCCGG ACTTCGTACA CGTCCTTTAC AAGGGGCGCA TCGTCAAGTC GGGAGGAGCC GAGCTGGCTC TCACGCTCGA AGAAAAAGGC TACGACTGGA TCAAGGAAGA GATAGGAGAA | 660 720 750 | | | | |
|------------|---|--|--|--|--|--|
| 5 | (2) INFORMATION FOR SEQ ID NO:146 | | | | | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1383 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | |
| 15 | (ii) HOLECULE TYPE: DNA (genomic) | | | | | |
| 10 | (iii) HYPOTHETICAL: NO | | | | | |
| | (iv) ANTI-SENSE: NO | | | | | |
| 20 | <pre>(vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS</pre> | | | | | |
| 25 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11383 | | | | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:146 | | | | | |
| 30 | ATGGCTAAGG AGAAAACGAT CTACGTCTGC CGTTCGTGCG GAACCAAATA CGCCAAATGG CAAGGCAACT GCAATGCCTG TGGAGAGTGG AACTGCATTG ATGAGGAGAA GGTGCCGGCA CGGCATCGG GCAAGCATGC AGCCAAAGAGT TTTATGCCTC GGGAGCAGGA CAACCGGCCA AGACTCTTAC AGGATGTGGA GTCCGGCGAT GAAGAGCGTA TTCGCCTCGG CGATGAAGAG TTCGACCGCG TACTGGTGG AGGAATTGTC AAAGGAGCAT TTGTCCTGCT TGGCGGCGAG | 60 120 180 240 300 | | | | |
| 35 | CCGGGAATCG GTAAGTCCAC GCTTATCCTC CAGACGGTGC TGCGTCTGCC GCAGTTGCGC ACGCTCTATG TGTCGGCGA AGAAAGTGCC CGACAACTGA AGATGCGCGC CGAACGACTG GGGAAAGCAA TGAATGGGTG CTACGTATAC TGCGAAACGA ATATACAGAC GATCACACC CGATCACCT GTGATAGACT CTATACAGAC GGTCTTACCC GAGGAAATGG AAAGCTCGGC CGGCAGCGTG GGCAGATCC GCGAATGTGC CGCCTTACTG CTCAAATACT GCAAGACTAC GGGTTATCCCC GTCATCGTCA TCGGACACAT CACCAAAAGAA | 360 420 480 540 600 660 | | | | |
| 40 | GGTAGCATAG CCGGACCGAA GGTGCTGGAG CATATAGTGG ATACGGTGCT TCTCTTCGAC 720 GGGGATAAGC ATCATCTCTA CCGGATACTC CGAGGACAGA AGAACCGCTA TGGCAGTACT 780 TCCGAGCTGG GGATATACGA GATGCGGCAG GACGGTCTGC GTGGCGTGGA GAATCCGAGC 840 GAACATCTCA TCACACGCAA TAGGGAAGAC CTCAGTGGCA TAGCCATAGC CGTAGCGATG 900 | | | | | |
| 45 | GAGGGCATTC GCCCGATACT CATCGAAGCG CAGGCTTTGG TCAGCTCGGC CATTTATGCC 960 AATCCGCAGC GTTCGGCCAC GGGCTTCGAT ATTCGGCGGA TGAACATGCT CTTAGCCGTA 1020 CTGGAGAAAC GTGCCGCTT CAAGCTCATA CAGAAGGATG TGTTTCTGAA CATTGCCGGA 1080 GGTATCAAAA TAGCCGATCC GGCTACGGAT CTGCCCGTTA TCTCGGCAGT GCTGGCGTCG 1140 AGTCTGGACA TCGTTATCCC GCCGGCCGTA TGCATGACGG GCGAGGTCGG ACTCTCCGGA 1200 | | | | | |
| 50 | GAGATACGTC CCGTGAGCCG CATCGAGCAG CGCATAACGG AAGCGCGTCG CATAGGGTTC AAAGAGATAT TGGTACCGGC CGATAATTTC CGGCAGGAGG ATGCCGGCCG CTTCGGTATT CGGCTCGTGC CGGTCAGAAA GGTGGAGGAA GCCTTCCGCC ATCTGTTCTC GAAAGGAAGA GAA | 1260 1320 1380 1383 | | | | |
| 55 | (2) INFORMATION FOR SEQ ID NO:147 | | | | | |
| 60 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 813 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | |
| | (ii) HOLECULE TYPE: DNA (genomic) | | | | | |
| 65 | (iii) HYPOTHETICAL: NO | | | | | |
| | (iv) ANTI-SENSE: NO | | | | | |
| 70 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | | | | | |
| | (ix) FEATURE: | | | | | |
| 7 5 | (A) NAME/KEY: misc_feature (B) LOCATION 1813 | | | | | |

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147
      ATGAATAGCA GACATCTGAC AATCACAATC ATTGCCGGCC TCTCCCTCTT TGTACTGACA
      TTGGGCGGCT GCTCCGTAGC CCAACAAGAT ACGCAGTGGA CTCTCGGCGG AAAGCTCTTT
                                                                                120
 5
      ACTTCGGCGT GGATACAACG TTCGGCCGAA TATCAAGCGC TTTGCATTCA GGCATACAAC
                                                                                180
      ATCGCTACGG AAAGAGTGGA CGCTCTACCG GCAGAACGTA AACAAGGAGA TAGGCCTTAT
                                                                                240
      GCCATCGTAA CGGACATAGA CGAAACCATT TTGGACAATA CGCCTAACTC CGTGTATCAG
                                                                                300
      GCTCTCAGGG GCAAGGATTA TGATGAAGAG ACTTGGGGGGA AATGGTGTGC ACAGGCCGAT
                                                                                360
      GCCGACACAC TGGCAGGAGC TTTGTCTTTC TTCCTCCATG CAGCGAACAA GGGGATCGAG
                                                                                420
10
      GTCTTTTACG TCACCAACCG CAGAGACAAT CTGCGCGAAG CAACTCTTCA GAACCTTCAG
                                                                                480
      CGTTACGGAT TCCCCTTTGC CGATGAAGAA CATTTGCTTA CGACCCATGG GCCATCCGAC
                                                                                540
      AAAGAACCCC GTCGGCTCAA AATACAAGAA CAGTATGAAA TAGTATTGCT CATAGGAGAC
      AACTTGGGCG ACTTCCACCA CTTCTTCAAT ACGAAAGAAG AGTCCGGACG CAAACAGGCT
                                                                                660
      CTGGGCCTGA CAGCCGGGGA GTTTGGCCGG CACTTCATCA TGCTGCCCAA TCCCAACTAC
                                                                                720
15
      GGATCTTGGG AACCGGCATG GTACGGCGGG AAGTATCCGC CACTGCCCGA AAGAGACAAA
                                                                                780
      GCACTTAAAC AACTGCACTC ACAGAACAGC AGA
                                                                                813
      (2) INFORMATION FOR SEQ ID NO:148
20
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1251 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
25
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
30
           (iv) AHTI-SEHSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
35
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1251
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:148
      ATGAGCACCA ATATAGATGT ACAACAGATC AAACAGCGTT TCGGCATCAT CGGTAGCAGT
                                                                                60
      CCGCTGATGG AACATGCCAT ACGAGTGGCA GCACAGGTGG CTCCTACCGA CATGTCCGTC
                                                                                120
      CTCGTGACGG GGGAGAGCGG TTCGGGGAAA GAGTTCTTCC CACAGATAAT CCACTACTAC
                                                                                180
45
      AGCGCCCGGA AACATCATAG CTACATTGCA GTCAATTGCG GAGCCATCCC CGAAGGAACC
                                                                                240
      ATCGATTCCG AGCTGTTCGG ACACCGCAAA GGTTCCTTTA CCGGAGCCGT ATCGGATCGC
                                                                                300
      AAGGGGTACT TCGAAGAAGC ATCCGGCGGC ACGATCTTTC TGGACGAAGT GGGCGAACTG
                                                                                360
      CCTTTGCCCA CGCAGGCGAG GCTGCTGAGG GTGCTGGAGA CGGGCGAGTT CATCCCCGTA
                                                                                420
      GGAGCCAGCC AGTCGCAGAA GACGGATGTC CGTATCGTAG CGGCGACGAA TGTGAACCTC
                                                                                480
50
      AAGGAGGCGG TAGCGAACGG GAAGTTCCGG GAAGACCTCT TCTTCCGGCT CAATACGGTA
                                                                                540
      CCGATCGAGG TGCCTGCGCT GCGTATGCGA CCGGACGACG TGCCCTTGCT TTTTCGCCGA
                                                                                600
      TTCGCCGCCG ACAGCGCCGA GAAGTATCGG ATGCCTCCGC TGCGCCTATC GGACGAAGCC
                                                                                660
      CGTACCATAT TAATGCGTTA CCGCTGGCCC GGCAATGTGC GAGAGCTGCG CAATATAACC
                                                                                720
      GACAGGCTGA GCATCCTGGA GGAGGAGCGG ACGGTATCGG CAGAGACCAT CACTCGCTAC
                                                                                780
55
      CTGGACGCTG AGGGGATGCA AGACCTCCAC CCCGTCGTGA TCCGACGGAA CGAAACGACC
                                                                                840
      GAAGCGGACA AACAATCCC CCATTACGAG CGCGAAATCA TCTACCAGGT GCTATACGAT
ATGAAGAAAG AGATAGCCGA TTTGAAGGGG ATGATGAACC GCCTGGCGCA CCACGAACAG
                                                                                900
                                                                                960
      CCCTCATGGC CTGTAGGGTC GGACGTCTGG GGCAACGACG ACAAGCGCAC CGCAGATCCG
                                                                              1020
      AAGTGGGGCG TCAGCACGCA CAAGGCCCCC ATCGCGAACG CGGCAGAACC CGTGGAGCCG
                                                                              1080
60
      ATACAGGAAG CCAGCGAATA CACCGAGGAT CCGGTTTCGC TGGAGGAGGT AGAGAAGAAA
                                                                              1140
      ATGATTTCCC TTGCATTGGA ACGCCACGGC GGAAGGCGCA AGCAGACAGC CGAGGAACTG
                                                                              1200
      AAGATTTCGG AGCGGACACT ATACCGTAAA ATCAAGGAGT ATGGACTGGA A
                                                                              1251
65
      (2) INFORMATION FOR SEQ ID NO:149
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1806 base pairs
                 (B) TYPE: nucleic acid
70
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
75
         (iii) HYPOTHETICAL: NO
```

```
(iv) AHTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
 5
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1806
10
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:149
      ATGATCCTAC TCTTCGGCGG TACTACGGAA GGCCGTGCCG CAGCTCGCGT GCTGGATGAA
      GCGGGAAGTC CGTTTTTCTA CTCCACCAAA GGCAATCTGC AAGAGATCCA GAGTAGCCAC
                                                                                  120
                                                                                  180
15
      GGCCATCGTC TGACAGGAGC CATGACGGTT GCCGACATGG TTTCGTTTTG TCGGAAAGAA
       GAGATCCGAC TGATCGTGGA CGCCGCTCAT CCTTTCGCCG AAGAATTGCA CGCTTCAGTG
                                                                                  240
                                                                                  300
      GCAGAAGCCT CTGAACAAC AGGTATCCCC GTAGTAAGAT ACGAGAGACA ATACCCTCCA
                                                                                  360
      CGCGAAGAAG GTATCGTCTG GTGTGCAAAC TACGATACGG CTGCCGAGCG GATGCTTGGC
      GATGGCGTGC AGCGTCTGCT GATGCTCACA GGAGTGAATA CGATCCCCAA GCTGGCTGCT
                                                                                  420
20
      TTCTGGAAAG AGCGCACCAC CTTTTGCCGC ATATTGAAGC GAGACGAATC GGTTGCTTTG
                                                                                   480
      GCAGAGAAGA ACGGCTTTCC TGCGGAGCGC ATCGTTTTCT TCGAACCGCA TGCGGACGAG
                                                                                  540
                                                                                  600
      GAGCTGATGC AAGCCGTTCG CCCCGATGCC ATTATCACAA AAGAAAGCGG AGAGAGCGGT
      TACTTCCGAG AAAAGATAGA AGCTGCCCGA CGGATGGGCA TCCGTATATA TGCCGTCGTA CGTCCCCCTT TGCCTCCTTC ATTCATTCCC GTAGGCGGGC CTGTCGGTTT GAGACGGGCG
                                                                                   660
25
                                                                                  780
       GTAGAACGCC TCGTGCCGGG ATTCTTTCA CTCCGAAGCG GATTCACTAC CGGCACCACA
       GCTACCGCTG CAGTAGTAGC AGCCATGTAC CGATTGATGG GGCTTGGCTC TCTCGCCGAA
                                                                                  840
                                                                                  900
       GCTCCCGTAG AATTGCCTTC GGGCGAAATA GTCAGTCTGC CCATAGCGGA AATTCGAGAG
       GAAGAAGATG CTGTCGTATC CGCAGTCCTG AAAGATGCAG GTGATGATCC GGATGTGACC
                                                                                  960
       AATGGCATGG CGGTATGCGC TACGATCAGG CTCAATCCCG AACATGAGGA AGTCCGCTTC
                                                                                 1020
30
       CTGCAGGTG AAGGGGTGGG GGTAGTGACG CTCCCCGGCC TCGGTCTGGA GGTCGGAGGT
                                                                                 1080
       COGGCTATCA ACCTOGTACO TOGACGAATG ATGACAGCAG AGGTACGCCG ACTCTATGCG
                                                                                 1140
       CAGGGAGGTG TGGATATTAC GATTAGCGTA CCCGAAGGCC GAGAGGCTGC TACCCAGACA
                                                                                 1200
                                                                                 1260
       TTCAATCCCC GACTCGGCAT ACGGGACGGC ATCTCTATTA TCGGAACATC GGGAGTCGTG
       AAACCTTTTT CGGCCGAAGC GTTCGTTGGT GCCATCCGTA AGCAAGTGGG TATTGCCACC
                                                                                 1320
35
       GCCTTGGGAG CCAATCATAT CGTCCTCAAT TCGGGAGCCA AGAGTGAGCG TTATGTAAAA
                                                                                 1380
       GGAGCCTATC CGGCACTCAT TCCACAGGCC TTTGTGCAGT ATGGCAATTT CGTCGGCGAA
TCACTCAGTT GTGTAGCTTC CTTCCCTTCT GTCCGTTCGG TAACGGTAGG AATCATGCTC
                                                                                 1440
                                                                                 1500
       GGCAAAGCAG TGAAACTCGC CGAAGGCTAT CTGGATACGC ACAGTAAAAA GGTAGTGATG
                                                                                 1560
       AATCGGGATT TCCTGCACGA ACTGGCTCGT CAGGCAGGTT GTTCGGAAGA CATCCATGCC
                                                                                 1620
40
                                                                                 1680
       ATAATAGACA GCCTGAATTT GGCTCGTGAG CTATGGACTA TGCCGAGTGC GGAGGACAGC
       GATCGACTGC TACGAAAGAT TGCCGAACGA TCTTGGGAAA CTTGCCGCCC ATCGGTACCA
                                                                                 1740
       TCGGCCGAAT TAGAACTCCT GCTGATCGAT GAGTCCGGAG CGATTCGTTT TCGTATCGGT
                                                                                  1800
                                                                                  1806
       GGAGAA
45
       (2) IMFORMATION FOR SEQ ID NO:150
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1329 base pairs
50
                  (B) TYPE: nucleic acid
                  (C) STRAHDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
55
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
60
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROHOMAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
65
                  (B) LOCATION 1...1329
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:150
                                                                                    60
       ATGTTGAGGA CTTTCCGAAT CGGTGGTATT CACCCCCCG AAAACAAGTT GTCGGCAGGC
70
       AAGCCCGTAG AGGTGTTGCC TATCCCCTCA CAGGTAGTCA TCCCTCTTGG TCAGCACATC
                                                                                   120
       GGTGCACCGG CAACTGCCAC GGTCAAGAAA GGGGATGAAG TTAAGGTCGG GACTATCATT
                                                                                   180
       GCTCAGGCCG GAGGATTCGT ATCAGCTAAT ATCCACTCAT CTGTGTCGGG TAAGGTGCTG
                                                                                   240
       AAGATCGATA ACGTATACGA CTCAAGCGGC TATCCCAAGC CCGCAGTCTT CATTAGCGTA GAAGGTGACG AATGGGAAGA GGGCATCGAT CGCTCACCAG CCATCGTCAA AGAATGCAAT
                                                                                   300
                                                                                   360
 75
       CTGGATGCAA AAGAAATCGT AGCCAAAATT TCTGCAGCCG GTATTGTGGG TCTTGGCGGT
                                                                                   420
```

```
GCTACCTTCC CTACCCATGT GAAGCTGTCC CCTCCTCCGG GCAACAAGC TGAGATCCTG
                                                                            480
      ATCATCAACG CCGTAGAGTG CGAGCCTTAT CTGACGAGCG ACCATGTCCT TATGCTGGAG
      CACGGCGAAG AGATCATGAT CGGCGTGAGT ATCCTGATGA AAGCCATTCA GGTAAACAAG
                                                                            600
      660
 5
      ACTGCATATC CGGGCATAGA GGTAATGCCG TTGAAGGTGC AATATCCTCA AGGCGGTGAG
                                                                            720
      AAGCAGCTGA TCGATGCAGT GATCCGCAAG CAGGTAAAAA GCGGTGCCTT GCCTATCAGC
      ACAGGTGCCG TAGTACAAAA CGTGGGTACG GTATTCGCCG TGTACGAAGC AGTACAGAAG
      AACAAGCCTC TGGTCGAGCG CATCGTGACG GTTACAGGAA AAAAACTGTC TCGTCCGTCT
      AACCTCCTCG TTCGTATAGG TACTCCTATT GCGGCTTTGA TCGAAGCAGC AGGTGGCTTG
                                                                            960
10
      CCGGAGAATA CGGGCAAGAT CATCGGCGGA GGTCCGATGA TGGGACGCGC TCTGCTGTCA
                                                                           1020
      CCGGATGTGC CTGTGACCAA AGGCAGCTCC GGAGTATTGA TTCTCGATAG AGAAGAGGCA
                                                                           1080
      GTTCGCAAGC CTATGCGCGA CTGTATCCGA TGCGCCAAGT GCGTCGGAGT GTGTCCGATG
                                                                           1140
      GGACTCAATC CGGCTTTCCT TATGCGCGAC ACCTTATATA AGAGCTGGGA AACAGCGGAA
                                                                           1200
      AAAGGCAACG TGGTTGACTG TATCGAATGC GGTTCGTGCA GCTTCACCTG TCCGGCCAAC
                                                                           1260
15
      CGTCCTCTGC TGGATTATAT CCGCCAAGCC AAGAAGACTG TGATGGGTAT CCAAAGAGCA
                                                                           1320
      CGTAAGCAA
                                                                           1329
      (2) INFORMATION FOR SEO ID NO:151
20
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1437 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
25
                (D) TOPOLOGY: circular
          (ii) HOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
30
          (iv) AHTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
35
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...1437
40
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:151
      ATGAAAAGAA TACAACTAAC TCTTATCGCT CTCTTCGCCG CTGTTGCCGG TTTGGTCGCT
      CAAAATGCTT ACGAGGGAGT AATTTCATAT AAAATTTCGT TGGACAAAAC CGGAAACAAG
                                                                            120
      GTTGTACTGA ATGGTGCGGC AGATATGAGT AATTTAAAGC TCAAGAGCAC TCAGATGATC
                                                                            180
45
      ATTGTTACGC CTATTCTTCG TTCAGAAGAT GGTACCAGCC GGGTGGAATT TCCTTCGGTA
                                                                            240
      GTCATTACAG GCCGCAATAG AACAAAAGCT CTCAAGCGTG AAATCGCATT TAGTTCGGCT
      TTGCCCCAAG CAAAACATGC AGCTCAATAC ATTCGCCGTC ATAATGGGAA GAGCGAGCAG
                                                                            360
      TTTGCTTTTA CAGGAGAACA TGCTTATGCA TCATGGATGA TGGATGCCAA GTTTGTGGTT
      CGTGAGGAGG TACGAGGTTG TGCTAAATGC CCTGTAGGTC TCTCGAGTAA TATTGTTCCT
                                                                            480
50
      TTTGATCCAC TCTTCAATCC GGCAGAGGCT CCTTATTTGT TGGCACACAT TACTCCGGCA
      GAAGAAGTGG AAAAACAGCG AGAGTCCAGC TTCGATGCTT ATATCAACTT CAAAGTCAAT
      AAGGCAGATG TCCTTCCTGA GTATCGCAAC AATAAGGCGG AGTTAGAGAA AATCAAAGAA
      TTTGTAAGCA CCGTTAAGGC TAATCCAAAC TATTCGGTCA ATAAAATGAT CATCGAAGGG
                                                                            720
      TTTGCTTCTC CCGAGGCTTC AATAGCCCAC AATAAGGCTT TGTCGGAGCG CCGTGCTAAA
                                                                            780
55
      AGACTCGCGG AAGAATTGGT GCGTAAGTAT GGCAAAACAT TGCCGAATAT AACCACTGAA
                                                                            840
      TTCGGCGGTG AAGATTGGAA GGGGCTGAAA CTGGCTATCG AAAAGAGTGA TATAGCCGAT
                                                                            900
      CGTGACCGCG TATTGGAGAT AATCAACTCC GATAAATATG CCGATGATGA TGCACGTGAA
                                                                            960
      CAGGCTCTGA AGCAACTTTC GTCTTATCGT TATATCTTGG ATCAGATCTA TCCGAATTTG
                                                                           1020
      CGTCGCAATA CGATAACCAT GGGGTATATC GTTCGTGATT ATACCCTCGA AGAAGCTCGT
                                                                           1080
60
      GAAATCATTA AGACTGCTCC GAAAGAACTT AGTGAGGCCG AAATGTACCG TGTGGCAATG
                                                                           1140
      TCTTATCCTG AGGGGCACCA AGAGCGTTTG TTTGCTCTGA ATACGACCCT TAAGTATTTC
                                                                           1200
      CCTGAAAGTG TAACGGGCCG AATCAATTTG GCTGTAGCCG CTTTTAATGG TGGAGACGTT
                                                                           1260
      CAACAGGCAA TTGCTCTGTT GAGTCCGATT CAGACAGAAA AGGGTGTAAG CAATATCCTT
                                                                           1320
      GGAGCTGCTT ATGCTCGTAC GGGAGATTTT GCTCGTGCCG AAACCTTCTT CCGTAAGGCC
                                                                           1380
65
      GTTGCAGAAG GAGATGCAAA TGCGCAGCGC AACCTCGATA TGCTGCTTGG CAAAAAG
                                                                           1437
      (2) INFORMATION FOR SEQ ID NO:152
70
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1149 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
75
```

WO 99/29870 PCT/AU98/01023

```
(ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
 5
          (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
10
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1149
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:152
15
      ATGGCAGAAA AAAGAGACTA TTACGAAGTC CTCGGTGTAT CGAAGAATGC CACCGACGAT
      GAACTGAAAA AAGCATATCG CAAGAAGGCT ATCCAATACC ATCCTGATAA GAACCCCGGT
                                                                                 120
                                                                                 180
      GACAAGGAGG CCGAAGAGCA CTTCAAAGAG GTAGCTGAAG CCTACGACGT ATTGAGCGAT
      CCGCAGAAGC GCAGTCAATA TGACCAGTTC GGCCATGCCG GATTGGGCGG AGCTGCCGGT
                                                                                 240
20
      GGAGGTTTCA GCGGAGGCGG TATGTCCATG GAGGATATTT TCAGTCGCTT CGGTGATCTA TTCGGTGGGT TCGGCGGTTT CGGCGGATTC TCCGATATGG GCGGTGGCAG TCGCAGACGT
                                                                                 300
                                                                                 360
      GTTCGCAGAG GGTCTGACCT GCGAGTACGA GTGAAGCTTT CTTTGGCCGA TATAAGTAAA
                                                                                 420
      GGTGTGGAGA AGAAAGTGAA GGTAAAAAAG CAGGTAGTGT GCAGCAAATG TCGTGGCGAT
                                                                                 480
      GGCACGGAAG AAGCCAATGG CAAGACTACC TGCCAGACCT GCCATGGAAC CGGCGTGGTT
                                                                                 540
25
      ACACGTGTGA GCAACACTTT CCTTGGGGCC ATGCAGACCC AGAGCACTTG TCCCACTTGC
                                                                                 600
      CACGGAGAG GTGAGATCAT CACGAAGCCA TGCTCCAAGT GTAAGGGCGA AGGTGTGGAG
                                                                                 660
      ATCGGCGAAG AGGTGATCTC ATTCCACATC CCTGCCGGTG TAGCCGAAGG AATGCAAATG
                                                                                 720
      TCCGTGAACG GCAAGGGAAA TGCCGCGCCC CGAGGAGGCG TGAATGGCGA CTTGATAGTC
                                                                                 780
      GTGATCGCCG AGGAACCGGA TCCGAATCTG ATCCGCAATG GCAACGATCT GATATACAAT
                                                                                 840
30
      CTGCTTATAT CCGTTCCGTT GGCTATAAAA GGAGGTAGTG TGGAAGTGCC GACGATAGAC
                                                                                 900
      GGACGAGCCA AGATCCGCAT CGAGGCGGGG ACACAACCCG GCAAGATGCT GCGTTTGCGC
                                                                                 960
      AATAAGGGGT TGCCCAGCGT AAACGGCTAT GGCATGGGAG ACCAACTGGT GAATGTCAAT
                                                                                1020
      GTCTATATCC CCGAATCGAT CGATGCCAAA GATGAGCAGG CTATCGCAGC GATGGAAAAC
                                                                                1080
                                                                                1140
      TCGGACAGCT TCAAACCTAC CGATGCTGCT CGTAAGGATA TAGACAAGAA ATACAGAGAG
35
                                                                                1149
      ATGCTGGAT
      (2) INFORMATION FOR SEQ ID NO:153
40
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 879 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
45
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
50
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
55
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...879
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:153
60
      ATGAAAAAC TGATTTTAGC GACTTTGGGA CTTATGGCCA TTGCCATGCT CTCATGTTCA
                                                                                  60
      AGCAACAACA AGGATTTGGA GAACAAAGGG GAGGCTACTC TTTTGGTAAC GTTTGGTAGC
                                                                                 120
      TCCTATAAAG CTCCACGCGA AACCTATGCG AAGATTGAGA AGACTTTTGC CGCAGCTTAT
                                                                                 180
      CCCGATCAAA GGATAAGCTG GACATACACG TCTTCTATTA TCCGAAAGAA ACTGGCTCAG
                                                                                 240
65
      CAGGGTATT ATATCGATGC TCCGGATGAG GCTTTGGAGA AATTGGCTCG TCTGGGTTAT
                                                                                 300
      AAGAAGATCA ATGTACAGAG TCTTCATGTG ATTCCCGGCC GAGAATATGA TGAGATGATC
                                                                                 360
      GACTTTGTCA ATAAGTTTAA GGCAGCACAT AGTGATATTA CTGTGAAGGT AGGGGCTCCG
                                                                                 420
      CTTTTCGATA CCGATGAAGA TATGCGCGAG GTGGCAGAGA TCTTGCACAA GCGTTTTCAG
                                                                                 480
      CAAACGATAG AGAAAGGTGA AGCTATTGTA TTCATGGGAC ACGGCACCGA GCATGCTGCC
                                                                                 540
70
      AATGACAGGT ATGCCCGTAT CAATAAGATC ATGAAGAACT ATAGCAAGTT CATGATCGTC GGAACCGTCG AGTCCGATCC CTCTATCAAT GATGTTATTG CCGAACTGAA AGAAACCGGT
                                                                                 660
      GCCACGGCCG TAACAATGAT GCCGCTGATG AGTGTGGCAG GCGACCATGC TACGAATGAT
                                                                                 720
      ATGGCCGGAG ATGAGGACGA TAGCTGGAAG ACGTTGCTGA CCAATGCCGG CTACACAGTT
                                                                                 780
      TCTATAGACA AGCTGGACAA TGGCAATTTC TCAGCTCTTG GAGATATAGA AGAGATCCGG
                                                                                 840
75
      AATATCTGGC TCAAGCATAT GAAAGCCACC TCTGCTCGC
                                                                                 879
```

```
(2) INFORMATION FOR SEQ ID NO:154
 5
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1068 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
10
           (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
15
           (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
20
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1068
          (xi) SEQUENCE DESCRIPTION: SEO ID NO:154
25
      ATGACATCCG TCAGCCACTT ACGTACAATT TCTGTCGCAG GTATCCTGGC TGCGCTGGGA
      GGGGCTGTAC TCATTCTCTT CGGGGTTAAT CTCTTCCTCG GCTCGGTGGC TATTCCGATG
                                                                               120
      AGCGAGATCT TCCGACATCT TTTTTCAGAT CGTCCCGAAG GAGGAGAAGC ACTCGTGCAC
                                                                               180
      TACAATATCC TATGGAAATC CCGCCTGCCC GAAGCCCTCA CGGCTGCTTT TGCCGGCGCA
                                                                               240
30
      GGTTTATCCG TTAGTGGCTT GCAGATGCAG ACCGTCTTTC GCAATCCTTT GGCCGGTCCG
                                                                               300
      TCCGTTCTCG GCATCAGCTC CGGTGCCAGT TTGGGTGTTG CTTTGGTCGT TCTGCTGAGC
                                                                               360
      GGCTCGCTGG GAGGAGTGGC ATTGAGTAGC CTGGGTTATA TGGGCGAGGT GGCCATGAAT
                                                                               420
      ATAGCCGCTG CCGTAGGCTC GCTGGCAGTA ATGGGGCTGA TCGTTTTTGT CAGCACCAAG
                                                                               480
      GTGCGCAGCC ACGTTACGCT GCTCATTATC GGCGTTATGA TCGGATATGT AGCCACTGCC
                                                                               540
35
      GTCATCGGGG TATTCAAGTT TTTCAGTATC GAAGAAGATA TTCGGGCATA CGTAATTTGG
                                                                               600
      GGGTTGGGCA GCTTTTCCCG TGCCACGGAT TCGCAACTGA GTTTCTTTGC CATTCTGATG
                                                                               660
      TTGATCTTTA TTCCGGCCGG TATGCTCCTT GTCAAGCAGT TGAATCTCTT ATTGCTGGGA
                                                                               720
      GAAAGCTACG CACGTAATCT GGGACTGAAT ACTCGTCGGG CACGGCTGCT CGTGATCTCT
                                                                               780
      TCCGCCGGTT TGCTCATCGC TACCGTCACG GCCTATTGCG GTCCCATCGG CTTTTTGGGG
                                                                               840
40
      ATGGCTGTGC CACACTTGGC ACGGGTTATC TTTCACACAT CGGATCATCG GATCCTGATG
                                                                               900
      CCTGCTACCT GTTTGATTGG AAGTGCTCTG GCTCTTTTCT GCAATATCAT TGCTCGTATG
                                                                               960
      CCGGGGGTTTG AGGGGGCTTT GCCCGTCAAT TCCGTAACGG CTTTGGTGGG AGCACCTATT ATCGTCACCG TTTTGTTCCG GCGCAGACGC TTCAAGGAAG AAACCGAC
                                                                              1020
                                                                              1068
45
      (2) INFORMATION FOR SEQ ID NO:155
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2271 base pairs
50
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
55
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
60
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
65
                 (B) LOCATION 1...2271
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:155
      ATGCGGACAA AAACTATCTT TTTTGCGATT ATCTCTTTTA TTGCTCTATT GTCGTCTTCT
                                                                                60
70
      CTGTCGGCTC AGAGCAAAGC CGTTTTAACC GGTAGTGTGT CGGATGCCGA AACCGGAGAG
                                                                               120
      CCTCTTGCCG GTGCTCGAAT CGAAGTCAAA CACACCAACA TAGTAGCCGG TGCCGATGCC
                                                                               180
      GGCGGACATT TCGAGATCAA GAACCTGCCG GCAGGGCAGC ATACTATTAT ATGTTCGTTG
                                                                               240
      GGGGGGTATG GACAGAAAGA GGAGGTGGTT GCCATCGAAG CCGGACAGAC CAAAACGATC
                                                                               300
      TCTTTTGCAT TGCGACTGCG AACGAACAAC TTGGAGGAAG TCGTCGTTAC CGGTACCGGT
                                                                               360
75
      ACACGTTACC GCTTGGTCGA TGCTCCTGTG GCAACGGAAG TCCTTACCGC TAAGGACATA
```

```
GCCTCTTTCT CGGCTCCTAC TTCCGAGGCC TTATTGCAGG GGCTGAGTCC GTCTTTTGAC
                                                                                  480
      TTCGGCCCCA ATCTGATGGG CTCTTTCATG CAGCTGAACG GCCTTAGCAG TAAGTATATC
      CTCATCCTTA TCGATGGTAA GCGTGTGTAC GGCGATGTAG GCGGTCAGGC CGATTTGAGT
                                                                                  600
      CGTATTTCTC CTGATCAGAT CGAACGGATC GAACTGGTGA AAGGTGCTTC GAGTTCGCTC
                                                                                  660
 5
      TACGGATCCG ATGCCATCGC CGGGGTAATC AATGTGATCA CAAAAAAGAA TACGAATCGA
                                                                                  720
      CTGAGTGCAT ATACGTCACA TCGCATATCG AAGTACAACG ATCGGCAAAC CAATACTTCG
                                                                                  780
      CTCGATATAA ACATCGGTAA GTTCAGTAGC AATACCAACT ATTTCTTCTA CCATACGGAT
                                                                                  840
      GGCTGGCAGA ATAGTCCGTT CGAAATAAAA AAGAAAAAAG GATCCGGCGA ACCGGTCTTG
                                                                                  900
      GAGGAAACGT ATAAGAAAAC TTTTCGTGCA CAGGAAAATC AGGGTGTAAG CCAATCGCTT
                                                                                  960
10
      TCCTATTATG CAACTAACAA TCTTAGCTTC AGCGGAAATG TGCAGTACAA TAAACGTCAG
                                                                                 1020
      ATCTTCACTC CGACTTTTTC CGAAAAGAAG GCCTATGACA TGGATTATCG TGCTTTGACG
                                                                                 1080
      GCTTCACTCG GTACGAACTA TCTTTTCCCC AATGGTCTGC ATACGCTTTC TTTCGATGCC
                                                                                 1140
      GTCTACGATC GCTTCCGTTT CGGATATTTG TATCATGACA AGGACAGCAG TGAGAGCCTG
                                                                                 1200
      ATCAACAACC AAGGTCAGAC CGAGCAACCC ACATTCTTTC CGGGTCAGCT ACGCAATAAA
                                                                                 1260
      AACGATCAGA TCCGATACAC GGCAGAGGCT CGCGGTGTAT TTACACTGCC TTATGCGCAG
15
                                                                                 1320
      AAACTGACCG GCGGTTTGGA GTATTTCCGT GAGGAATTGA TCTCTCCCTA TAATTTGATT
                                                                                 1380
      ACCGACAAGG CAGATGCTTC CACGCTCTCT GCTTATGTAC AAGATGAATG GAAACCGCTC
                                                                                 1440
      GATTGGTTCA ATATGACAGC CGGTTTCCGT CTGGTACACC ATCAGGAGTT CGGTACACGA ATGACGCCTA AGGTATCCAT ACTCGCCAAG TATGGGCCGC TGAACTTCCG CGCTACGTAT
                                                                                 1500
                                                                                 1560
20
      GCTAACGGCT ATAAGACTCC CACGCTGAAA GAGCTTTTTG CACGGAACGA ACTCACCACT
                                                                                 1620
      ATGGGTTCGC ACAATCTCTA TCTCGGCAAT GCGGATCTTA AGCCACAGAT GTCGGATTAT
                                                                                 1680
      TATGCTTTGG GCTTGGAGTA CAATCAAGGC CCTATCTCGT TCAGTGCAAC GGTTTATGAC
                                                                                 1740
      AATGAACTTC GCAATCTGAT CTCCTTTATG GATATACCGA CCTCACCCGA GCACGAAGCT
                                                                                 1800
      CAGGGAATCA AGAAAACCAA GCAGTATGCC AACATAGGAA AAGCTCGCAG CCGCGGCCTT
                                                                                 1860
25
      GATGTCCTAT GTGATGCCTC TATCGGTTGG GGTATCAAGT TAGGAGCCGG ATACAGCCTC
                                                                                 1920
      GTGGAAGCTA AGAATCTCCA GACGGATGAG TGGCTGGAAG GAGCTGCACG TCATCGTGCC
                                                                                 1980
      AATGTGCACG CCGATTGGT TCACTACTGG GGTCAGTATA GACTTGGCGT GAGCCTTTTC
      GGCCGTATTC AGAGCGAGCG TTACTACAAA GACGGCAATG CTCCGGACTA TACCTTGTGG
                                                                                 2100
      CGACTOGCCA CATCGCATCG TTTCGCTCAT TTCCGCCACA TCATCCTGGA TGGAACGCTC
                                                                                 2160
30
      GGTATAGACA ACCTGTTTGA CTACGTGGAT GATCGTCCTA TGGGTGTCAA TTATGCTACC
                                                                                 2220
      GTAACGCCGG GACGTACTTT CTTTGCTCAA ATAGCGATTC GATTCAACAA C
                                                                                 2271
       (2) INFORMATION FOR SEQ ID NO:156
35
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 993 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
40
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
45
           (iv) AUTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
50
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...993
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:156
      ATGACGGACA ACAAACAACG TAATATCGTA TTCCCGGCGT TTCTCCTCTT GCTGGGAGTC ATCGCAGTGG TGACGATCGT TGGTTTTTC ATGCTCAGAC CGGCCGAGGA GATTATCCAA
                                                                                  120
                                                                                  180
      GGACAGATAG AAGTGACCGA ATACCGAGTG TCCAGCAAAG TGCCCGGGCG CATCAAGGAA
60
      CTTAGGGTAT CCGAGGGACA GCAGGTGCAG GCCGGCGATA CCCTCGCTGT CATCGAAGCC
                                                                                  240
      CCCGACGTAG CGGCTAAGAT GGAGCAGGCA AAGGCTGCCG AAGCAGCTGC ACAGGCTCAG
                                                                                  300
      AACGCCAAGG CTCTCAAAGG AGCACGCAGC GAACAGATAC AGGCAGCCTA TGAGATGTGG
                                                                                  360
      CAGAAAGCTC AGGCCGGCGT AGCCATAGCG ACCAAGACAC ACCAGCGCGT GCAGAACCTC
                                                                                  420
      TATGACCAGG GAGTGGTACC GGCTCAGAAG TTGGACGAAG CCACTGCCCA GCGCGATGCG
                                                                                  480
65
      GCCATCGCTA CGCAAAAAGC GGCCGAAGCC CAGTACAATA TGGCTCGCAA CGGTGCCGAA
                                                                                  540
      CGCGAAGACA AGCTGGCAGC TTCTGCCCTC GTCGATAGAG CGAGAGGAGC CGTCGCCGAG
                                                                                  600
      GTGGAGTCGT ACATCAACGA AACCTACCTC ATCGCCCCAC GGGCAGGCGA AGTGTCGGAG
                                                                                  660
      ATATTCCCCA AAGCCGGCGA ACTCGTAGGT ACCGGCGCAC CTATCATGAA TATCGCCGAG
ATGGGCGATA TGTGGGCCAG CTTTGCCGTT CGTGAGGATT TCCTCAGCAG CATGACCATG
                                                                                  720
                                                                                  780
70
      GGAGCCGTTC TGGAGACTGT GGTGCCGGCT CTGAATGAAG AAAAAGTACG CTTCAAGATC
                                                                                  840
      ACATTCATCA AGAACATGGG TACCTATGCT GCCTGGAAAG CGACCAAGAC AACAGGGCAG
                                                                                  900
      TACGACCTGA AGACCTTCGA GGTAAAGGCC ACCCTTGCGG ATAAAGACAA GGCACAAAAG
                                                                                  960
      CTACGCCCGG GTATGTCCGT GATCATACGC AAG
                                                                                  993
75
```

```
(2) INFORMATION FOR SEQ ID NO:157
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 801 base pairs
 5
                 (B) TYPE: nucleic acid
                 (C) STRAHDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
10
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
15
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
20
                 (B) LOCATION 1...801
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:157
      ATGCGTATTG TCAGTAATTT TTTGTTCGTC TCTTTTTCGG TTTTGCTTTT TGCATCATGC
25
      CGTTCCCAGC GAGAAAAGGT CGTTTACCTG CAAGATATCC AAACTTTTAA TCGGGAGATT
      ATCGCTAAAC CATATGACGT AAAAATTGAG AAGGACGATG TGCTGAACAT CCTTGTCAGC
                                                                                180
      AGTAGAGACC CGGAGCTTTC AACGCCCTAC AACCAAGTGT TGACCACTCG TGCACTGGCC
                                                                                240
      CGCAACGGCT ATGGAACGAA CTCGAACGAA GGCTTCCTGG TCGATTCGAA AGGGTACATC
                                                                                300
      AATTATCCTA TTTTAGGCCA GATCTATGTA GAGGGCCTTA CTCGTACCGA ACTGGAGAAG
                                                                                360
30
      GAGATACAGA AGAGGATTAT TTCCAGTGGA TTTATCAAGG ATCCTACGGT AACGGTGCAG
                                                                                420
      CTTCAAAATT TCAAGGTGTC GGTTTTGGGA GAGGTGAATC ATCCGGGTTC GATGTCGGTA
                                                                                480
      AAAGGAGAGC GAATAACTCT TTTGGAAGCG ATCGGAATGG CCGGAGACCT GACAATCTAT GGTCGCCGCG ATCGGGTTTT TGTGATTAGA GAAACCGATG GGCATCGCGA GGTTTTCCAG
                                                                                540
                                                                                600
      ACGGATCTCA GAAAGGCCGA CTTGCTCGCA AGCCCCGTGT ACTATCTGCA TCAGAACGAC
                                                                                660
35
      GTCATCTATG TGGAGCCGAA CGACAAGAAA ACACAGATGA GCGAGATCAA CCAGAATAAT
                                                                                720
      AACGTAAACG TATGGCTGAG TGTTACCTCC ACTTTGGTAT CCATTTCCAC GCTGACGATT
                                                                                780
      ACGATAATAG ATAAGACCAA A
                                                                                801
40
      (2) INFORMATION FOR SEQ ID NO:158
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1707 base pairs
                 (B) TYPE: nucleic acid
45
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
50
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
55
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1707
60
          (xi) SEQUENCE DESCRIPTION: SEO ID NO:158
      ATGAAAAAGA CCAATCTGTT TTTATCTCTG CTGGTGATCT TTATCACCGG TAGTTTTATG
      ACTGCCTGTG CACAGAAGTC CAAGACGAAC AAACTCACCG AAGAAGATCG GAGCCGCAAT
                                                                                120
65
      GAGTATGTAC AGTCGATGGA TGTGCTTAGC AATATTATCG GTAACGTCAG GCTGTATTTC
                                                                                180
      GTCGATACCA TAAGTATCAA ACATATGACT CGGCGTGGTA TAGATGCGAT GTTGGGCGGG
                                                                                240
      CTTGACCCCT ATACCGAATA CATTCCTTAC GAGGAAATGG ATGAACTGAA ATTGATGACT
                                                                                300
      ACGGGAGAGT ATGCCGGAGT CGGAGCTATC ATATCGCAGC GCCCGGATAG TGCTGTGATT
                                                                                360
      ATCCAGAGAC CTATGGAAGG TATGCCCGCA GACGAAGCAG GATTGATAGC AGGCGACCGC
70
      ATCCTGACTA TCGATGGGAA AGACTTCCGC AAATCCACCA CACCGAAAGT AAGCCAAGCA
                                                                                480
      CTGAAAGGGA TAGCCGGTAC TGTTGCAAAG GTGACAGTAA TGCGCTATGG CGAAACCAAA
                                                                                540
      CCTCGTACTT TTTCCGTGAA ACGTCAAAAA GTGATTATGA ATTCCGTCAC TTACAGCGGA
                                                                                600
      ATGCTCGATG GCTCGATAGG ATATATCCGC TTGAACAACT TTACGGACAA AAGTGCAGAA
                                                                                660
      GAGGTGCGCA CGGCCTTGTT GGATCTTCGT GACAAACAAG GAGCGAAAGG TCTCATTTTG
                                                                                720
75
      GATTTAAGAG GCAATGGTGG CGGACTGATG CAGGCTGCTA TCGAGATAGT CAATCTGTTC
                                                                                780
```

PCT/AU98/01023

| 5 10 | GTCCCTAAGG GCAAAGAGGT GGTAACGACC AAAGGTCGCA TTGCAGAGTC GGCGTCCG TTTCGCACAT TGACTGAACC GATCGACAG AAACTCCCGA TAGTAGTCCT GATCGATG CAATCGGCAT CTTCCTCGGA GATTGTAGCC GGAGCACTGC AGGATATGGA CAGGGCTG CTGATGGGAC AAAAGAGCTA TGGCAAAGGG CTTGTACAAA CGACTCGTCA GCTACCAT AACGGCGTGA TCAAAATTGAC TACGGCCAAG TACTACATC CAAGCGGACG TTGTATTC CACAAAATCT TTTACACTGC TGCCGGAAGA CGTTTAGGAAA ATCCTGC GACTCGAGG TCAAACAAGA TACAGCTGCG ACATTACTT ATTATATGGC CATCAATA GACGTTTTCG ATTTCGTCAC AGGTTATGTA AAACGATTCC CAAGCCGGA GATTTTTCCA TAACGAACGA GGACTATGCA GCTTTCTGCA AGATGAGACT TTTGACTATG TCCCCGGAAGC CAACTCGGAC CTTAAAGCAAC TGGCGAAAA CTACGCGAAAAACAACAC TGCCCGAAGA CTTGACAAAC TACGCCGAAAAACAACAC TGGCCAAGAC CAACTCGGAC CTTAAAGCAAC TACGCGAAAAA CAAACTATCT CAACAAAAAAAAAA | GA 900 TA 960 AC 1020 AG 1080 TG 1140 CT 1260 AG 1320 AA 1380 TA 1440 CC 1500 |
|------------|---|--|
| 15 | ATTGTCACTC GCTATTATTA TGAGCGAGGC AGTATCCGCC AGAGTTTGCC GGAAGATA GTAGTCAAAG AAGCTATTAA GCTGCTGAAG GACCATCCGG AACAAATTCG ACAGATCC GCAGCTCCGA AAGCAGAGAA TAAAGGG | AG 1620 |
| 20 | (2) INFORMATION FOR SEQ ID NO:159 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2943 base pairs (B) TYPE: nucleic acid | |
| 25 | (C) STRAHDEDNESS: double (D) TOPOLOGY: circular | |
| | | |
| | (ii) MOLECULE TYPE: DMA (genomic) | |
| 20 | (iii) HYPOTHETICAL: NO | |
| 30 | (iv) ANTI-SENSE: NO | |
| | | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| 35 | (II) OKSINIZIN TOKITIMKONONING GINGTVIBID | |
| | <pre>(ix) FEATURE: (A) NAME/KEY: misc feature</pre> | |
| | (B) LOCATION 12943 | |
| 40 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:159 | |
| 10 | (AI) DECORNOR DESCRIPTION. SEQ ID NO. 133 | |
| | ATGCAAAACA AAGGATTTGT GATTGTTATC ACATCGGCTC TGGCCATCAT CTGTGCGT TACCTGTCAT TCTCTTTCGT TACGAACCGT TACGAAAAGA AGGCTAAGGC GATGGGCG | |
| | GTTGCCGGAA TGGCCTATCT TGATTCCATG TCGAATGAGA AGGTCTGGTT CGGCTACA | |
| 4 5 | CTGAAAGAAG CTCAAGCCCA GCAAATTGGT CTTGGCCTTG ACTTAAAGGG GGGTATGA | |
| | GTTATCTTGA AACTTAACGC AAGCGATCTG CTTCGTAACC TCTCTAACAA AAGTTTGG CCCAACTTCA ACAAAGCTCT GGAGAATGCT GCCAAGAGCA CGGAGCAATC CGACTTCA' | |
| | GATATTTCG TGAAGGAATA TCGCAAGCTC GATCCCAACG GTCGCTTGGC CGTTATCT | |
| F 0 | GGTTCGGGTG ACCTTCGCGA CCAGATTACC GCAAAGTCTA CGGATGCAGA CGTAGTGC | |
| 50 | CTGCTCAAAG AAAAATATAA TAGTGCTGTA GAAGCTTCGT TCAATGTGCT CCGTGCTC | |
| | ATCGATGCTT TCGGTGTGGT TGCACCTAAT TTGCAGCGAT TGGAAGGGCA AGGGCGTA' CTTGTCGAAC TCCCCGGAGT GAAAGACCCT GAGCGTGTTC GTACCCTTTT GCAACGCA | |
| | GCCAACCTAC AGTTCTGGCG TACATACAAA TTCGAAGAGG TCAGCGGAGA CTTGATCG | |
| 55 | GCCAATGATC GTCTGAGCGA ATTGGCTATG AACAACACGG ATGCTACCCC GGAAACAG | |
| 33 | | |
| | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGTACA AGCTGTAG | CT 840 |
| | | CT 840 FC 900 |
| | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGTACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAG ATGGCTCAGA TATCTGAAAT GCTCCAGCAA GCTCACGATC TGAAGGTTAC ACGTGAAGG | CT 840 FC 900 AT 960 AT 1020 |
| 6 0 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGTACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCCAGCAA GCTCACGATC TGAAGGTTAC ACGTGAAGG GTGCTTTTCC TCTGGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAA GGAGACCG | CT 840 FC 900 AT 960 AT 1020 AC 1080 |
| 6 0 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGTACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCAGGAAA GCTCACGATC TGAAGGTTAC ACGTGAAGA GTGCTTTTCC TCTGGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAA GGAGACCG CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGA | CT 840 IC 900 AT 960 AT 1020 AC 1080 AT 1140 |
| 6 0 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGTACA AGCTGTAGG GATTCTGCTTA CTGCAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCGTGAA TCGTGGCGGT GCAGTAGTG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCCAGCAA GCTCACGATC TGAAGGTTAC ACGTGAAGG GTGCTTTTCC TCTGGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAA GGAGACCGG CTCTACCGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGG GTAGTGACTT CCGCCAAGAG TGAATACCAA AATGACTTCG GTCGTTCCGA ACCGATCG TCGATGACGA TGAATGAAGA AGGTGCTCGT AAATGGGCGC GTATCACAAA GGATAACG | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 PG 1260 |
| 60 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCGTGAA TCGTGGCGGT GCAGTAGTG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCCAGCAA GCTCACGATC TGAAGGTTAC ACGTGAAGA GTGCTTTTCC TCTGGGTGC TAAAGCAATC GAAGACCCCG AACCAAAAA GGAGACCGA CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGGA ATCCTGATTT GGGAGGTGA GTAGTGACTA CGCCAAGAG TGAATTCCAA AATGACTTCG GTCGTTCCGA ACCGATCGC TCGATGACGA TGAATGAAGA AGGTGCTCGT AAATGGGCGC GTATCACAAA GGATAACG GGACGGCCAA TCGCTATCGT TTTGGATGGT GTGGTTTATT CTGCTCCGAA CGTGAATGA | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 PG 1260 AT 1320 |
| 60 65 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGTACA AGCTGTAGG GATTCTGCTTA CTGCAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCGTGAA TCGTGGCGGT GCAGTAGTG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCCAGCAA GCTCACGATC TGAAGGTTAC ACGTGAAGG GTGCTTTTCC TCTGGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAA GGAGACCGG CTCTACCGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGG GTAGTGACTT CCGCCAAGAG TGAATACCAA AATGACTTCG GTCGTTCCGA ACCGATCG TCGATGACGA TGAATGAAGA AGGTGCTCGT AAATGGGCGC GTATCACAAA GGATAACG | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 PG 1260 AT 1320 AC 1380 |
| | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCTT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCACGAAC GCTCACGATC TGAAGGTTAC ACGTGAAGA GTGCTTTTCC TCTGGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAAA GGAGACCCC CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGA GTAGTGACCA TGAATGAAGA AGGTGCTCGT AAATGACCT GTCGTTCCGA ACCGATCGT TCGATGACCA TGAATGAAGA AGGTGCTCGT AAATGGCGC GTACTACAAAA GGATAACGT GAGATCACGG GCGTCGCTC TCAGATCTCC GGGCACTTCA CCCGTGGAGGA CGTGAATGA GAGATCACGG TCCTCAACTC CGGTAAAATG GATGCTCACGG TAAGCACGGAACGT CTTGCCAACG TACTCAACTC CGGTAAAATG GATGCTCACGG GATTCTTGCC GTCCTCGCC CTTGCCAACG TACTCAACTC CGGTAAAATG GATGCTCACGG GATTCTTGCC GTTCCTGCC CTTGCCAACG TACTCAACTC CGGTAAAATG GATGCTCACGG GATTCTTGCC GTTCCTGCCC CTTCCACCC CTACGCTGG TGCCCACTCC CCGTGGAGGA GCCGGTGG CTTCGCCAACG TACTCAACTC CGGTAAAATG GATGCTCACGG GATTCTTGCC GTTCCTGCC CTTCTCACCC CCCTGGAGGA GCCCGCGGACTCCA CCCGTGGAGGA GCCGGTGG CTTCTCAACTC CGGTAAAATG GATGCTCCCG GATTCTTCC GTTCCTCCCC CTTCTCACCC CCCTGGAGGA GCCCGCTGATCACCC CCGTGGAGGA GCCGGTGG CTTCTCACCC CCCTCGCGCC CACCCC CCCTGGAGGA CGCCGCTGC CTTCCTCCCCC CCCCCCC CCCCCCC CCCCCCCCCC | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 PG 1260 AT 1320 AC 1380 AC 1440 PC 1500 |
| | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCGTGAA TCGTGGCGGT GCAGTAGTG GTGTGGCTC TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCACGATC TGAAGGTTAC ACGTGAAG GTGCTTTTCC TCTGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAA GGAGACCG CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGAG ATCCTGATTT GGGAGGTGC TCGATGACGA TGAATGAAGA AGGTGCTCGT AAATGGCCC GTATCACAAA GGATAACG GGACGGCAA TCGCTATCGT TTTGGATGGT GTGGTTTATT CTGCTCCGAA CGTGAATG GAGATCACG GCGGTCGTC TCAGACTCC GGGCACTTCA CCGTGGAGGA GCCGGTG CTTGCCAACG TCTCACGCT CGGTAAAATG GATGCTACGG GATCCTGCTCC CTTGCCAACG TCTCACGCTGG TGCCGAACG GATCCTCC CGGTACACG GATCCTCC GGGCACTTCA CCGTGAGAGA GTGATTGGTC CTACGCTGGG TGCCGACCC ATTAAAGCCG GATTCTTGCC GGCTCTTACGC TTCCTGCCCGA CGTGATCTC GTGCTACGG TGCCTACGG TGCCGACCC ATTAAAGCCG GATTCTTGCC GGCTCTTACGCC GTTCTTTGCC GGCTCTTACGCC GTTCTTTCCC GGGTCCTTACCCCCTTCCCCCAC GTTCTTCCCCGCCCCTTCACCCCCCCCCC | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 PG 1260 AT 1320 AC 1380 AC 1440 PC 1500 PC 1560 |
| | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCTT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCCAGCAA GCTCACGATC TGAAGGTTAC ACGTGAACG CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGGG ATCCTGATTT GGGAGGTG; CTCATGCAAC TCTATGCTAT TCGTACCAA AATGACTTCG GTCGTTCCGA ACCGATCG* TCGATGACGA TGATACCAA AGGTCTCT CGGCTACAAA GGAGACCG; GGACGGGCAA TCGCTATCGT TTTGGATGTG GTGGTTTATT CTGCTTCCGAA CGTGAATG; GAGATCACGG GCGGTCGCTC TCAGATCTC GGGCACTTCA CCGTGGAGGA GGCCGGTG; CTTGCCAACG TACTCAACTC CGGTAAAATG GATGCTACG TACCCAACA ACGGAACG GTGATTGGTC TACCCTGGG TGCCGATCC TTCACATTCG GTTCTTGCC GTTCCTCCCGACGCTC CTTGCCTACGC TACCCAGCTC TTCACACTC GTTCTTTCCCC GCTTTTGGTTA TCCTGATTGT AAACAGCTTC TTCACATTGG GTTTCTTGCC GGGTCTTTCCCC GCAAACGGCG CATTGATTGT AAACAGCTTC TTCACATTGG GCGTATTGGC TTCTTTCCCC | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 AC 1380 AC 1340 PC 1500 PC 1560 AT 1620 |
| | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCT TCTCTGCTTA CTCCGTGAA TCGTGGCGGT GCAGTAGTG GTGTGGCTC TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCACGATC TGAAGGTTAC ACGTGAAG GTGCTTTTCC TCTGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAA GGAGACCG CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGAG ATCCTGATTT GGGAGGTGC TCGATGACGA TGAATGAAGA AGGTGCTCGT AAATGGCCC GTATCACAAA GGATAACG GGACGGCAA TCGCTATCGT TTTGGATGGT GTGGTTTATT CTGCTCCGAA CGTGAATG GAGATCACG GCGGTCGTC TCAGACTCC GGGCACTTCA CCGTGGAGGA GCCGGTG CTTGCCAACG TCTCACGCT CGGTAAAATG GATGCTACGG GATCCTGCTCC CTTGCCAACG TCTCACGCTGG TGCCGAACG GATCCTCC CGGTACACG GATCCTCC GGGCACTTCA CCGTGAGAGA GTGATTGGTC CTACGCTGGG TGCCGACCC ATTAAAGCCG GATTCTTGCC GGCTCTTACGC TTCCTGCCCGA CGTGATCTC GTGCTACGG TGCCTACGG TGCCGACCC ATTAAAGCCG GATTCTTGCC GGCTCTTACGCC GTTCTTTGCC GGCTCTTACGCC GTTCTTTCCC GGGTCCTTACCCCCTTCCCCCAC GTTCTTCCCCGCCCCTTCACCCCCCCCCC | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 AC 1380 AC 1340 PC 1560 AT 1660 AT 1680 |
| 65 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTGTACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCTCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAAAGACTCCCGTCAAAAAAA GGAGACCCCG AAACCAAAAAA GGAGACCCCC CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGACTCTCTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGACTCTCTCTACGAAC TCAAAGACAACAAAAA GGAGACCCCG AAACCAAAAAA GGAGACCCCG AAACCAAAAAA GGAGGACCCCG CTCTACGAAC TCGCCAAGAG TGATTCCAAA AATGACTTCG GTCGTTCCGA ACCGATCCCCGACACGACA | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 AC 1320 AC 1380 AC 1440 PC 1560 AT 1620 AT 1680 PT 1740 CG 1800 |
| 65 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTACA AGCTGTAGG GATTCTGCTA CTGCAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCTT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCAGCAAA GCTCACGATC TGAAGGTTAC ACGTGAAGA GTGCTTTTCC TCTGGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAAA GGAGACCCC CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGA GTAGTGACGA TGAATGAAGA AGGTGCTCGT AAATGACGT GTCGTCCCGAA CGTACCGAAC GGACGGCAA TCGCTATCGT TCTGGATGCT AAATGAGCC CCGTGGAGGA GGCCGGTGA GAGATCACGG GCGTCGTC TCAGATCTC GGGCACTTCA CCGTGGAGGA GGCCGGTGA GTGATTGGTC CTACCCAACC TACCCAACCA TGCCGAGCCC AAACCAACAAA GGATAACG GTGATTGGTC CTACCCAACCA TCGCTACGT TCAGATCTC GGGCACTTCA CCGTGGAGGA GGCCGGTGG GCTTTGGTTA TCCTGATGTG TACCCAACTC TCAGATCTC GGCCACTCA GGTTATCTCC GCCAACGGC CATTGATTGT TACCAACTC TCACATTGG GCGTATTGCC GGGTCTTACG GCCAACGTAC TTATCTCGG GCGTATCAAA GAAGAGCTTC TCCCCAACG TACCCAACGT TACCCAACGT TCCACATTGG GCGTATTGCC GGGTCTTACG GCCAACGTAC TTATCTCCGG GCGTATCAAA GAAGAGCTTC TCCACATTGG GCGTATTGCC TTCTTTCCC GCCTGCGTTA CGGATGGTTA TGGCAACGCT TTCTCTCCCA TCTTCGACTC GAACGTTACACCATC CCGGTATCAT TTCTCTGCCA TCTTCTCCCA TCTTCGACTC GAACGTTTACACTTCT TCCACATTCG GGCCGATTCC CGGTTTTGCCACCACCACCACCACCACCACCACCACCACCACCACCA | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 AC 1320 AC 1380 AC 1440 PC 1560 AT 1620 AT 1660 AT 1680 PT 1740 PC 1860 |
| 65 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTACA AGCTGTAGG GATTCTGCTA CTGTAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCTT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCCAGCAA GCTCACGATC TGAAGGTTAC ACGTGAACG CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGGG ATCCTGATTT GGGAGGTG GTAGTGACTT CCGCCAAGAG TGATATCCAA AATGACTTCG GTCGTTCCGA ACCGATCGT TCGATGACGA TGATATCCAA AATGACTTCG GTCGTTCCGAA GGATAACGC GGACGGGCAA TCGCTATCGT TTTGGATGGT GTGGTTATT CTGCTACGAA GGATAACGC GGAGTCACGG GCGGTCGCT TCAGATTCC GGGCACTTCA CCGTGGAGGA GCCGGTGG CTTGCCAACG TACTCAACTC CGGTAAAATG GATGCTACG TTCTTGCCAACG GCTTTGGTTA TCCTGATTGT TTCACATTGC GTTCTTGCC GCTTTGGTTA TCCTGATTGT TACACTGT TTCACATTGG GCTTATTGCC GGGTCTTC GCCAACGTCA CATTCACACT TTCACATTGC GCTTACGC GGGTCTTC GCCAACGTCA CCCTCTCGGG TATCCAAA GAAGAGCTTC TTCACATTGC GCTTATTGCC GGGTCTTC GCCAACGTCA CCCTCTCGGG TATCCAAA GAAGAGCTTC GCCGTGATTGC GCCAACGTCA CCCTCTCGGG TACCCAACGTTC TTCACATTGC GCCTGGTAT GCCTGTCGC GCCAACGTAC TATCTCAA GAAGAGCTT TTCACATTGC GTCCTGCGACCGACCGTAC ACTATTATTA CCGGTTACT TACCGCTTC TACGGGACGG GGCCGATTCC GAACGTTAC ACTACTTGA TTATCGGTCT TATCGCTTCT TACGGGACGG GCCCGATTCC GAACGTTAC ACTACTTGA TTATCGGTCT TATCGCTTCT TACGGGACGG CGTGTTTCTT GACTCGTACCGACCGTTACACCGACCGTTCC CGGTTTTCCCGACCGACCCGAC | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 AC 1380 AC 1380 PC 1560 PC 1560 AT 1620 AT 1680 PT 1740 PC 1860 PC 1860 PC 1860 PC 1920 |
| 65 | CCTGCAACTA CTGACTCTGT AGCTGCAACA GCCGATTCTG CTGCTACA AGCTGTAGG GATTCTGCTA CTGCAGCACA AAAAGAGGCC AAGGATGCTA CTCGTAAAGA CGCACTCTT TCTCTGCTTA CTCCCGTGAA TCGTGGCGGT GCAGTAGTGG GTGTGGCTCG TCGTGCTAA ATGGCTCAGA TATCTGAAAT GCTCAGCAAA GCTCACGATC TGAAGGTTAC ACGTGAAGA GTGCTTTTCC TCTGGGGTGC TAAAGCAATC GAAGACCCCG AAACCAAAAAA GGAGACCCC CTCTACGAAC TCTATGCTAT TCGTACCAAT CGTACGGGAG ATCCTGATTT GGGAGGTGA GTAGTGACGA TGAATGAAGA AGGTGCTCGT AAATGACGT GTCGTCCCGAA CGTACCGAAC GGACGGCAA TCGCTATCGT TCTGGATGCT AAATGAGCC CCGTGGAGGA GGCCGGTGA GAGATCACGG GCGTCGTC TCAGATCTC GGGCACTTCA CCGTGGAGGA GGCCGGTGA GTGATTGGTC CTACCCAACC TACCCAACCA TGCCGAGCCC AAACCAACAAA GGATAACG GTGATTGGTC CTACCCAACCA TCGCTACGT TCAGATCTC GGGCACTTCA CCGTGGAGGA GGCCGGTGG GCTTTGGTTA TCCTGATGTG TACCCAACTC TCAGATCTC GGCCACTCA GGTTATCTCC GCCAACGGC CATTGATTGT TACCAACTC TCACATTGG GCGTATTGCC GGGTCTTACG GCCAACGTAC TTATCTCGG GCGTATCAAA GAAGAGCTTC TCCCCAACG TACCCAACGT TACCCAACGT TCCACATTGG GCGTATTGCC GGGTCTTACG GCCAACGTAC TTATCTCCGG GCGTATCAAA GAAGAGCTTC TCCACATTGG GCGTATTGCC TTCTTTCCC GCCTGCGTTA CGGATGGTTA TGGCAACGCT TTCTCTCCCA TCTTCGACTC GAACGTTACACCATC CCGGTATCAT TTCTCTGCCA TCTTCTCCCA TCTTCGACTC GAACGTTTACACTTCT TCCACATTCG GGCCGATTCC CGGTTTTGCCACCACCACCACCACCACCACCACCACCACCACCACCA | CT 840 PC 900 AT 960 AT 1020 AC 1080 AT 1140 PT 1200 AC 1380 AC 1340 PC 1500 PC 1560 AT 1680 PT 1740 PC 1860 PC 1980 PC 1980 PC 1980 |

```
ATCATTCCGG TGATTATCAT CGTTTTGGGA CTTATAGCTT CATTTACAAT CGGTCTCAAT
      AGGGGTATTG AATTCTCCGG AGGACGTAAC TACGTAGTTA AATTCGACCA GCCTGTATCT
                                                                             2160
      TCCGAAGCCG TTCGTTCGGC CTTGTCTTCT CCCCTGCAGG AAAAGGTATT GGTTACCTCC
                                                                             2220
      ATCGGTACTG AAGGGACAGA GGTGCGTATA TCTACGAACT ATAAGATCCA GGAGGAAAGC
                                                                             2280
 5
      GAAGAAACTG AAGCAGAGAT TACTGACAAA TTGTATCAGA GCCTGAAAGG TTTCTACACC
                                                                             2340
      CAGCAGCCTA CTGCTGATCA GTTCTTGGAC AATATCATTA GCTCTCAGAA AGTAAGTCCC
                                                                             2400
      AGTATGTCGA GTGACATCAC GAGAGGTGCT ATTTGGGCTG TGCTGTTATC GATGATCTTC
                                                                             2460
      ATGGCCATTT ATATTCTGAT TCGCTTCCGT GACATTTCTT TCTCTGCCGG GGTATTCGTA
                                                                             2520
      TCTGTGGCCG CTACTACATT CTGCATTATT GCTCTGTATG CGTTGCTGTG GAAGATTCTG
                                                                             2580
10
      CCCTTCACCA TGGAGATCGA TCAGAACTTC ATCGCTGCTA TTCTGGCTAT CATCGGTTAC
                                                                             2640
      TCGCTCAATG ACACCGTGGT TGTATTTGAC CGTATCCGAG AGACGATGAA ATTGTACCCC
                                                                             2700
      AACAGAGATC GCTATCAGGT GATCAACGAT GCCCTTAATT CAACATTGGG TCGAACATTA
                                                                             2760
      AATACGTCTT TGACTACGTT TATCGTTATG TTGGTAATCT TCATCTTTGG AGGTGCTACG
                                                                             2820
      ATGCGTAGTT TCACGTTCTC GATCCTGCTC GGTATCGTTA TCGGTACATA CTCTACGCTC
                                                                             2880
15
      TTTGTTGCTA CACCCCTTGC CTACGAGATC CAAAAGCGCA AGCTCAACAA AGCAGCTAAG
                                                                             2940
      AAA
                                                                             2943
      (2) INFORMATION FOR SEO ID NO:160
20
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 3051 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
25
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
30
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
35
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
                (B) LOCATION 1...3051
40
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:160
      ATGAAAAGAA TGACGCTATT CTTCCTTTGC TTGCTGACGA GCATTGGGTG GGCTATGGCC
      CAGAATAGAA CCGTGAAGGG TACAGTTATC TCCTCCGAGG ATAATGAGCC CCTGATCGGC
                                                                             120
      GCGAATGTCG TGGTTGTCGG AAACACCACT ATCGGTGCTG CAACCGACTT GGATGGCAAC
                                                                             180
45
      TTCACGCTTA GCGTGCCTGC CAATGCCAAA ATGTTGAGAG TGTCCTATTC CGGTATGACT
                                                                             240
      ACCAAAGAG TCGCCATCGC TAATGTGATG AAGATCGTAC TGGATCCGGA CTCTAAGGTT
      CTGGAGCAGG TAGTTGTATT GGGTTACGGT ACGGGACAGA AACTCAGCAC TGTTTCCGGT
                                                                              360
      TCTGTGGCCA AAGTGTCCAG CGAAAAGCTC GCGGAAAAGC CCGTTGCCAA TATCATGGAT
                                                                             420
      GCCCTCCAAG GTCAGGTAGC CGGTATGCAG GTTATGACTA CATCCGGTGA CCCTACTGCC
                                                                             480
50
      GTCGCTTCTG TGGAGATCCA TGGTACAGGG TCGTTGGGGG CAAGCTCTGC ACCATTGTAT
                                                                             540
      ATCGTGGATG GTATGCAAAC TTCTTTGGAT GTTGTGGCTA CGATGAATCC GAATGATTTT
                                                                             600
      GAATCTATGT CCGTTTTGAA AGATGCTTCT GCAACATCTA TTTATGGAGC TCGTGCTGCA
                                                                             660
      AACGGAGTCG TTTTCATTCA AACGAAGAAA GGTAAAATGA GCGAGAGAGG TCGTATTACC
                                                                              720
      TTTAATGCCA GTTACGGGAT TTCTCAAATC CTGAATACTA AGCCCCTTGA TAATATGATG
                                                                             780
55
      ACTGGAGATG AATTGCTGGA TTTTCAGGTG AAGGCAGGTT TTTGGGGGAA CAATCAAACC
                                                                             840
      GTTCAGAAGG TTAAAGATAT GATCCTTGCC GGAGCTGAAG ATTTGTATGG CAATTATGAT
                                                                             900
      TCTTTGAAAG ATGAGTATGG TAAGACATTG TTCCCAGTGG ATTTTAATCA TGATGCAGAC
                                                                             960
      TGGCTCAAGG CTTTGTTTAA AACAGCACCC ACCAGTCAAG GTGATATTTC TTTCTCCGGA
                                                                            1020
      GGGTCTCAGG GAACTTCATA TTATGCCTCT ATAGGCTACT TCGATCAGGA AGGTATGGCT
                                                                            1080
60
      CGTGAACCGG CAAATTTTAA GCGCTATAGT GGCCGGCTCA ACTTCGAAAG TCGTATCAAT
                                                                            1140
      GAATGGCTGA AAGTTGGTGC AAATTTGTCT GGTGCGATAG CGAATAGACG ATCTGCCGAC
                                                                            1200
      TATTTTGGAA AGTATTATAT GGGGTCAGGT ACTTTCGGTG TGTTAACGAT GCCTCGTTAT
                                                                            1260
      TATAACCCTT TTGATGTGAA TGGGGATTTA GCAGATGTCT ATTACATGTA TGGAGCTACC
                                                                            1320
      AGACCTTCTA TGACAGAACC GTACTTCGCA AAAATGAGAC CGTTCAGTTC CGAATCACAT
                                                                            1380
65
      CAGGCCAATG TAAATGGTTT CGCCCAGATT ACTCCGATCA AAGGCCTTAC TTTAAAGGCA
                                                                            1440
      CAGGCTGGTG TTGATATTAC TAATACTCGC ACTTCTTCTA AGAGAATGCC CAATAATCCG
                                                                            1500
      TATGATTCTA CTCCTCTTGG GGAAAGAAGA GAAAGAGCTT ATCGAGATGT TAGCAAGTCT
                                                                            1560
      TTTACAAATA CGGCTGAATA TAAGTTTTCA ATTGATGAAA AACATGATCT TACAGCATTG
                                                                            1620
      ATGGGGCATG AATATTGA ATATGAAGGG GATGTTATTG GGGCATCTTC TAAAGGATTT
                                                                            1680
70
      GAAAGTGATA AGTTGATGTT ACTGAGCCAG GGAAAAACCG GAAATAGTTT GTCTTTGCCT
                                                                            1740
      GAACACAGAG TCGCTGAATA TGCCTATTTG TCTTTCTTTA GTCGTTTTAA TTACGGTTTT
                                                                            1800
      GACAAATGGA TGTATATAGA TTTCTCTGTT CGTAATGACC AATCCTCTCG ATTCGGATCC
                                                                            1860
      AATAATAGAA GCGCGTGGTT CTATTCTGTC GGTGGAATGT TTGACATATA TAATAAATTC
                                                                            1920
      ATTCAAGAAA GTAATTGGCT CAGTGATCTT CGACTGAAAA TGAGTTATGG TACAACGGGT
                                                                            1980
75
      AACTCGGAGA TTGGTAATTA CAACCACCAA GCACTCGTTA CTGTGAACAA TTATACTGAA
                                                                            2040
```

```
GATGCTATGG GGCTTAGCAT TTCTACAGCA GGCAATCCCG ACCTCTCGTG GGAAAAGCAG
                                                                                2100
      TCTCAGTTCA ACTTCGGTTT GGCTGCAGGG GCTTTCAATA ATCGCTTATC TGCAGAGGTA
                                                                                2160
      GATTTCTATG TCCGCACTAC GAATGATATG TTGATTGATG TCCCGATGCC TTATATCAGT
                                                                                2220
      GGTTTCTTCT CACAGTATCA GAATGTAGGC TCTATGAAAA ATACGGGTGT AGACCTTTCT
 5
      CTTAAGGGGA CGATCTACCA AAATAAGGAC TGGAATGTAT ATGCTTCTGC GAATTTCAAC
                                                                                2340
      TACAATAGAC AGGAAATAAC AAAGCTTTTC TTCGGTCTCA ATAAGTACAT GTTGCCTAAT
                                                                                2400
      ACCGGTACTA TATGGGAAAT TGGGTACCCC AATTCGTTCT ATATGCTGA ATATGCTGGA
                                                                                2460
      ATCGACAAAA AAACCGGTAA GCAGTTGTGG TATGTTCCTG GTCAAGTCGA TGCGGATGGT
                                                                                2520
      AATAAAGTTA CAACAAGCCA GTACTCAGCT GACTTGGAGA CACGAATTGA TAAGTCTGTT
                                                                                2580
      ACTOCTOCTA TTACAGGTGG TTTCTCCTTA GGTGCTTCTT GGAAAGGACT TTCTTTAGAT
10
                                                                                2640
      GCTGATTTTG CCTACATCGT TGGTAAATGG ATGATCAATA ATGACCGTTA CTTTACAGAG
                                                                                2700
      AATGCAGGTG GATTGATGCA ATTAAATAAA GATAAAATGC TATTGAATGC CTGGACAGAG
                                                                                2760
      GATAATAAAG AAACAGATGT TCCAAAATTG GGACAGTCTC CTCAGTTTGA TACGCATTTG
                                                                               2820
      TTGGAGAATG CTTCTTTCCT GCGTTTGAAG AATCTCAAAC TCACCTATGT ACTCCCCAAT
                                                                                2880
15
      AGTCTTTTTG CTGGGCAGAA TGTGATTGGT GGAGCTCGTG TCTATTTGAT GGCGCGCAAT
                                                                                2940
      CTGTTAACTG TTACGAAGTA TAAAGGCTTT GACCCTGAAG CAGGGGGGAA TGTGGGAAAA
                                                                                3000
      AATCAATATC CTAATTCTAA GCAGTACGTT GCGGGTATTC AGTTGTCTTT C
20
      (2) INFORMATION FOR SEQ ID NO:161
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 3042 base pairs
                 (B) TYPE: nucleic acid
25
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
30
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
35
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...3042
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:161
      ATGACGCTAT TCTTCCTTTG CTTGCTGACG AGCATTGGGT GGGCTATGGC CCAGAATAGA
      ACCGTGAAGG GTACAGTTAT CTCCTCCGAG GATAATGAGC CCCTGATCGG CGCGAATGTC
                                                                                120
45
      GTGGTTGTCG GAAACACCAC TATCGGTGCT GCAACCGACT TGGATGGCAA CTTCACGCTT
                                                                                180
      AGCGTGCCTG CCAATGCCAA AATGTTGAGA GTGTCCTATT CCGGTATGAC TACCAAAGAG
                                                                                240
      GTCGCCATCG CTAATGTGAT GAAGATCGTA CTGGATCCGG ACTCTAAGGT TCTGGAGCAG
                                                                                300
      GTAGTTGTAT TGGGTTACGG TACGGGACAG AAACTCAGCA CTGTTTCCGG TTCTGTGGCC
                                                                                360
      AAAGTGTCCA GCGAAAAGCT CGCGGAAAAG CCCGTTGCCA ATATCATGGA TGCCCTCCAA
                                                                                420
50
      GGTCAGGTAG CCGGTATGCA GGTTATGACT ACATCCGGTG ACCCTACTGC CGTCGCTTCT
                                                                                480
      GTGGAGATCC ATGGTACAGG GTCGTTGGGG GCAAGCTCTG CACCATTGTA TATCGTGGAT
                                                                                540
      GGTATGCAAA CTTCTTTGGA TGTTGTGGCT ACGATGAATC CGAATGATTT TGAATCTATG TCCGTTTTGA AAGATGCTTC TGCAACATCT ATTTATGGAG CTCGTGCTGC AAACGGAGTC
                                                                                600
                                                                                660
      GTTTTCATTC AAACGAAGAA AGGTAAAATG AGCGAGAGAG GTCGTATTAC CTTTAATGCC
                                                                                720
55
      AGTTACGGGA TTTCTCAAAT CCTGAATACT AAGCCCCTTG ATAATATGAT GACTGGAGAT
                                                                                780
      GAATTGCTGG ATTTTCAGGT GAAGGCAGGT TTTTGGGGGA ACAATCAAAC CGTTCAGAAG
                                                                                840
      GTTAAAGATA TGATCCTTGC CGGAGCTGAA GATTTGTATG GCAATTATGA TTCTTTGAAA
                                                                                900
      GATGAGTATG GTAAGACATT GTTCCCAGTG GATTTTAATC ATGATGCAGA CTGGCTCAAG
                                                                                960
      GCTTTGTTTA AAACAGCACC CACCAGTCAA GGTGATATTT CTTTCTCCGG AGGGTCTCAG
                                                                               1020
60
      GGAACTTCAT ATTATGCCTC TATAGGCTAC TTCGATCAGG AAGGTATGGC TCGTGAACCG
                                                                               1080
      GCAAATTTTA AGCGCTATAG TGGCCGGCTC AACTTCGAAA GTCGTATCAA TGAATGGCTG
AAAGTTGGTG CAAATTTGTC TGGTGCGATA GCGAATAGAC GATCTGCCGA CTATTTTGGA
                                                                               1140
                                                                               1200
      AASTATTATA TGGGGTCAGG TACTTTCGGT GTGTTAACGA TGCCTCGTTA TTATAACCCT
                                                                               1260
      TTTGATGTGA ATGGGGGATTT AGCAGATGTC TATTACATGT ATGGAGCTAC CAGACCTTCT
                                                                               1320
65
      ATGACAGAAC CGTACTTCGC AAAAATGAGA CCGTTCAGTT CCGAATCACA TCAGGCCAAT
                                                                               1380
      GTAAATGGTT TCGCCCAGAT TACTCCGATC AAAGGCCTTA CTTTAAAGGC ACAGGCTGGT
                                                                               1440
      GTTGATATTA CTAATACTCG CACTTCTTCT AAGAGAATGC CCAATAATCC GTATGATTCT
                                                                               1500
      ACTCCTCTTG GGGAAAGAAG AGAAAGAGCT TATCGAGATG TTAGCAAGTC TTTTACAAAT
                                                                               1560
      ACGGCTGAAT ATAAGTTTTC AATTGATGAA AAACATGATC TTACAGCATT GATGGGGCAT
70
      GAATATATTG AATATGAAGG GGATGTTATT GGGGCATCTT CTAAAGGATT TGAAAGTGAT
      AAGTTGATGT TACTGAGCCA GGGAAAAACC GGAAATAGTT TGTCTTTGCC TGAACACAGA
                                                                               1740
      GTCGCTGAAT ATGCCTATTT GTCTTTCTTT AGTCGTTTTA ATTACGGTTT TGACAAATGG
                                                                               1800
      ATGTATATAG ATTTCTCTGT TCGTAATGAC CAATCCTCTC GATTCGGATC CAATAATAGA
                                                                               1860
      AGCGCGTGGT TCTATTCTGT CGGTGGAATG TTTGACATAT ATAATAAATT CATTCAAGAA
                                                                               1920
75
      AGTAATTGGC TCAGTGATCT TCGACTGAAA ATGAGTTATG GTACAACGGG TAACTCGGAG
                                                                               1980
```

```
ATTGGTAATT ACAACCACCA AGCACTCGTT ACTGTGAACA ATTATACTGA AGATGCTATG
      GGGCTTAGCA TTTCTACAGC AGGCAATCCC GACCTCTCGT GGGAAAAGCA GTCTCAGTTC
                                                                              2100
      AACTTCGGTT TGGCTGCAGG GGCTTTCAAT AATCGCTTAT CTGCAGAGGT AGATTTCTAT
                                                                              2160
      GTCCGCACTA CGAATGATAT GTTGATTGAT GTCCCGATGC CITATATCAG TGGTTTCTTC
                                                                              2220
      TCACAGTATC AGAATGTAGG CTCTATGAAA AATACGGGTG TAGACCTTTC TCTTAAGGGG
                                                                              2280
      ACGATCTACC AAAATAAGGA CTGGAATGTA TATGCTTCTG CGAATTTCAA CTACAATAGA
                                                                              2340
      CAGGAAATAA CAAAGCTTTT CTTCGGTCTC AATAAGTACA TGTTGCCTAA TACCGGTACT
                                                                              2400
      ATATGGGAAA TTGGGTACCC CAATTCGTTC TATATGGCTG AATATGCTGG AATCGACAAA
                                                                              2460
      AAAACCGGTA AGCAGTTGTG GTATGTTCCT GGTCAAGTCG ATGCGGATGG TAATAAAGTT
                                                                              2520
10
      ACAACAAGCC AGTACTCAGC TGACTTGGAG ACACGAATTG ATAAGTCTGT TACTCCTCCT
                                                                              2580
      ATTACAGGTG GTTTCTCCTT AGGTGCTTCT TGGAAAGGAC TTTCTTTAGA TGCTGATTTT
                                                                              2640
      GCCTACATCG TTGGTAAATG GATGATCAAT AATGACCGTT ACTTTACAGA GAATGCAGGT
                                                                              2700
      GGATTGATGC AATTAAATAA AGATAAAATG CTATTGAATG CCTGGACAGA GGATAATAAA GAAACAGATG TTCCAAAATT GGGACAGTCT CCTCAGTTTG ATACGCATTT GTTGGAGAAT
                                                                              2760
                                                                              2820
15
      GCTTCTTTCC TGCGTTTGAA GAATCTCAAA CTCACCTATG TACTCCCCAA TAGTCTTTTT
                                                                              2880
      GCTGGGCAGA ATGTGATTGG TGGAGCTCGT GTCTATTTGA TGGCGCGCAA TCTGTTAACT
                                                                              2940
                                                                              3000
      GTTACGAAGT ATAAAGGCTT TGACCCTGAA GCAGGGGGGA ATGTGGGAAA AAATCAATAT
      CCTAATTCTA AGCAGTACGT TGCGGGTATT CAGTTGTCTT TC
                                                                              3042
20
      (2) INFORMATION FOR SEQ ID NO:162
            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 2463 base pairs
25
                 (B) TYPE: nucleic acid
                    STRANDEDNESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
30
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
35
          (vi) ORIGINAL SOURCE:
                (A) ORGANISH: PORYPHYROHONAS GINGIVALIS
          (iz) FEATURE:
                (A) HAME/KEY: misc_feature
40
                (B) LOCATION 1...2463
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:162
      ATGAAGAAAA AGAATTTTTT GCTTCTTGGC ATTTTCGTTG CTTTGCTGAC TTTCATCGGC
                                                                                60
45
      AGCATGCAGG CACAACAGGC CAAAGATTAT TTCAACTTTG ACGAACGGGG CGAGGCCTAC
                                                                               120
      TTCTCATTCA AAGTGCCTGA TAGGGCCGTT CTACAAGAGC TGGCTCTGAT CATGTCCATC
                                                                               180
      GACGAGTTTG ACCCCGTAAC CAATGAAGCC ATTGCCTATG CCAGCGAAGA GGAGTTCGAG
                                                                               249
      GCATTCCTGC GCTATGGGCT CAAGCCTACA TTCTTGACTC CTCCATCCAT GCAGCGCGCT
                                                                               300
      GTCGAGATGT TCGACTACCG CTCAGGAGAA AAATACGAAT GGAATGCTTA CCCCACCTAT
                                                                               360
50
      GAAGCCTATA TCAGCATGAT GGAAGAGTTC CAAACAAAGT ATCCATCACT TTGTACTACT
                                                                               420
      TOOGTCATTG GCAAGTCOGT AAAGGATCGT AAACTGATGA TTTGCAAGCT GACGTCCTCT
                                                                               480
      GCCAATACAG GGAAAAAGCC TCGCGTGCTC TATACTTCTA CGATGCACGG AGACGAAACG
                                                                               540
      ACCGGATATG TGGTACTGCT CCGACTCATA GACCATCTGC TGTCGAACTA CGAATCCGAT
                                                                               600
      CCGAGGATTA AGAACATTCT GGATAAAACG GAAGTATGGA TCTGCCCTTT GACCAATCCG
                                                                               660
55
      GACGGAGCAT ACAGAGCCGG AAACCACACC GTACAAGGAG CTACTCGCTA CAATGCCAAC
                                                                               720
      AATGTCGATT TGAACCGTAA CTTCAAGGAT GATGTAGCCG GTGATCACCC CGATGGAAAA
                                                                               780
      CCTTGGCAGC CGGAGGCAAC TGCATTCATG GATTTGGAAG GAAACACCTC TTTCGTGCTC
                                                                               840
      GGTGCCAATA TACATGGAGG AACAGAGGTG GTGAACTATC CATGGGATAA TAAAAAAGAA
                                                                               900
      AGACATGCAG ACGATGAGTG GTACAAACTG ATCAGTCGCA ACTACGCAGC CGCTTGTCAG
                                                                               960
60
      AGTATTTCCG CCAGCTACAT GACCTCCGAA ACCAATTCGG GAATCATCAA CGGTTCAGAC
      TGGTATGTAA TTCGCGGAAG TCGTCAGGAC AATGCAAATT ATTTCCATCG TCTGCGAGAA
                                                                              1080
      ATTACCCTTG AAATCAGCAA CACGAAGTTG GTGCCGGCCT CTCAACTTCC AAAGTATTGG
                                                                              1140
      AATCTGAACA AAGAATCTCT GCTTGCTCTG ATCGAAGAAT CCTTATACGG CATCCATGGT
                                                                              1200
      ACAGTGACTT CCGCTGCGAA CGGACAGCCT CTCAAATGCC AGATCTTGAT AGAAAACCAT
                                                                              1260
65
      GACAAGCGCA ACTCCGATGT TTACTCCGAT GCTACCACAG GCTACTACGT ACGTCCTATC
                                                                              1320
      AAAGCCGGCA CTTATACGGT GAAATACAAA GCCGAGGGTT ATCCTGAGGC AACTCGTACC
                                                                              1380
      ATTACGATCA AGGACAAAGA AACCGTCATC ATGGACATTG CATTGGGCAA CTCGGTTCCT
                                                                              1440
      CTGCCTGTAC CCGATTTCAC AGCTTCTCCT ATGACCATCT CAGTAGGCGA AAGCGTCCAA
                                                                              1500
      TTCCAAGATC AAACGACAA TAACCCCACG AATTGGGAGT GGACGTTCGA AGGCGGACAG
                                                                              1560
70
      CCTGCCATGA GTACAGAGCA GAATCCGCTC GTATCCTATA GTCATCCCGG TCAGTACGAC
                                                                              1620
      GTTACGCTCA AAGTGTGGAA TGCAAGTGGT TCCAACACGA TTACGAAAGA AAAATTCATC
                                                                              1680
      ACTGTCAATG CCGTTATGCC TGTAGCTGAA TTCGTCGGTA CCCCGACGGA AATAGAAGAG
                                                                              1740
      GGCCAGACGG TATCTTTCCA AAACCAATCC ACCAATGCCA CCAACTACGT ATGGATATTC
                                                                              1800
      GATGGCGCA CTCCCGCTAC CAGTGAAGAC GAAAACCCGA CTGTGCTTTA CAGCAAAGCC
                                                                              1860
75
      GSCCAATACG ATGTCACGCT CAAGGCGATC AGTGCTTCCG GTGAAACGGT GAAGACGAAA
                                                                              1920
```

WO 99/29870 PCT/AU98/01023

```
GAAAAATACA TCACTGTCAA GAAAGCTCCG GTCCCTGCTC CGGTAGCCGA CTTCGAAGGA
                                                                            1980
      ACACCTCGAA AAGTAAAGAA AGGCGAGACA GTTACTTTCA AAGACTTGTC TACGAACAAT
                                                                            2040
                                                                             2100
      CCGACTTCAT GGCTTTGGGT GTTCGAAGGC GGCTCTCCTG CCACCAGCAC GGAGCAAAAC
      CCGGTGGTCA CCTACAATGA AACAGGCAAG TACGATGTCC AGCTGACTGC CACCAACGAG
                                                                            2160
 5
      GGCGGAAGCA ATGTGAAGAA AGCAGAAGAC TACATTGAGG TTATCCTCGA TGACAGTGTC
                                                                            2220
      GAGGACATAG TGGCACAGAC GGGTATCGTC ATTCGTCCGC AAAACGGAAC GAAGCAGATC
                                                                            2280
      CTCATAGAAG CCAACGCTGC TATCAAAGCG ATCGTTCTCT ATGACATCAA TGGACGGGTC
                                                                            2340
      GTACTCAAAA CTACTCCGAA TCAGCTCCGC TCGACCGTAG ATCTTTCCAT CCTGCCCGAA
                                                                            2400
                                                                            2460
      GGAATCTACA CCATCAATAT CAAAACGGAA AAATCCGCTC GCACGGAAAA GATCCATATC
10
      GGG
                                                                            2463
      (2) INFORMATION FOR SEQ ID NO:163
15
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 228 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
20
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
25
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
30
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...228
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:163
35
      ATGCGACTGA TCAAGGCTTT TCTCGTGCAA CTCTTACTGC TCCCCATTTT CTTCTACAAG
                                                                              60
      CGGTTTATAT CGCCGCTTAC ACCGCCTTCA TGCCGGTTTA CCCCCTCATG TTCGTCCTAT
                                                                             120
      GCCATCGAAG CCTTACGTAA ATATGGCCCG GGCAAAGGAC TATTGCTGAG CATCAAGCGT
                                                                             180
      ATTCTCCGCT GTCACCCGTG GGGTGGAAGT GGCTATGACC CCGTTCCG
                                                                             228
40
      (2) INFORMATION FOR SEQ ID NO:164
           (i) SEQUENCE CHARACTERISTICS:
45
                (A) LENGTH: 2577 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
50
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
55
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
60
                (A) NAME/KEY: misc_feature
                (B) LOCATION 1...2577
          (xi) SEQUENCE DESCRIPTION: SEO ID NO:164
65
      ATGGCATACG ACTITACACA AACATTCCGC AACAGCCTGG AGTACAGCTA TCAGGAAGCA
                                                                              60
      ACCOGTOTOG GOGTOGTAGO CGTGACGCAA GATATGCTCG TACTCGGTAT CATTCGCGAC
                                                                              120
      GGAGACAATG GCGCGATCGA CATCATGCGG CACTATGGGA TCAACTTGTA CGAACTCAAA
      CGGTTGATCG AGTTGGAAGC CATCGCCGAG AGTTTGCCTG CTTCGCCTGA GGGATCGCCC
                                                                              240
      ATCTTCACCC CTTCGGCTCG GGAGGCTATC GATGATGCCA CAGACATCTG TGCCGACATG
                                                                              300
70
      GAGGACGAGG CCGTCAGCCC GGTCCATCTG TTGCTGAGTA TCCTCAACTC GACACAGGAG
                                                                              360
      AGCTTAGTAC AAAAGATATT TATGAAACAA GGTATAAAAT ACGACACCAT CCTGTCGGAT
                                                                              420
      TACTTCGGAC AGCGCAACCC CTCCGAAGGG AAGTCTCCCT CCGAAATGGA GATCCTCGAC
                                                                              480
      GGGTACCAAG ACAACGACTT CGACGACGAA GAGGACGAAT CCTCTCCGCC TTCCGGGAAT
                                                                              540
      AGCGGGACAG GCGGAGGCTC CGGCGACGCC CCCGAACAGA ATACCGGCGG AGGCGATACT
                                                                              600
75
      ACCACCACGA CACGGAGTGG AGGCGACACG CCTGCACTGG ACACCTTCGG CACCGACATC
                                                                              660
```

PCT/AU98/01023

```
ACTGCCATGG CGGCAGCAGG CAAGCTCGAC CCGGTAGTGG GTCGGGAGCA GGAGATCGAA
                                                                               720
      AGGGTGATAC AGATACTCAG CCGCGCAAA AAGAACAATC CGGTGCTCAT CGGCGAACCC
                                                                               780
      GGTGTAGGCA AGAGTGCCAT CGTGGAAGGA CTGGCCGAAC GCATCGTGAA CAGGAAGGTG
                                                                               840
      AGCCGTATTC TTTTCGACAA GCGGATCATC AGCCTCGATT TGGCTCAGAT GGTAGCCGGC
 5
      ACCAAATATC GCGGACAGTT CGAAGAGCGG TTGAAAGCCG TGCTCGATGA GCTGAAGAAG
                                                                               960
      AATCCGCAGA TCATCCTCTT CATCGACGAG ATACATACCA TCGTGGGAGC AGGCTCTGCA
      GCCGGATCGA TGGATACGGC CAATATGCTC AAACCCGCTC TTGCCCGTGG ACAGGTACAG
                                                                              1080
      TGCATCGGAG CCACTACGCT GGATGAGTAT CGTAAGAACA TAGAAAAGGA CGGAGCACTC
                                                                              1140
      GAACGCCGCT TCCAGAAGGT GCCGATAGCC CCCTCGACTG CAGAAGAAAC GCTGACCATC
                                                                              1200
10
      CTGCAAAACA TCAAAGAGAA ATACGAGGAC TATCACGGTG TACGCTATAC GGACGAAGCG
                                                                              1260
      ATCAAAGCGG CAGTGGAACT GACCGATCGC TATGTATCCG ATCGTTTCTT CCCAGATAAG
                                                                              1320
      GCGATAGATG CCATGGACGA GGCCGGCGCG AGCGTCCATA TCACCAATGT GGTGGCTCCG
                                                                              1380
      AAAGAAATCG AGATACTGGA GGCCGAATTG GCATCGGTGC GAGAGAACAA GCTCTCGGCC GTAAAGGCTC AGAACTACGA ACTGGCTGCC TCCTTCCGCG ATCAGGAGCG GCGCACTCAG
                                                                              1440
                                                                              1500
15
      CAGCAGATAG CGGAAGAGAA GAAAAAATGG GAAGAGCAGA TGTCCAAGCA CCGCGAGACG
                                                                              1560
      GTGGACGAGA ATGTAGTGGC GCATGTAGTG GCGTTGATGA CAGGCGTTCC GGCTGAGCGG
                                                                              1620
      CTGAGCACGG GCGAAGGCGA ACGTCTGCGC ACGATGGCAG ATGATCTCAA GACCAAAGTA
                                                                              1680
      GTAGGTCAGG ACACAGCCAT CGAAAAGATG GTGCATGCCA TCCAGCGCAA TCGTCTGGGA
                                                                              1740
      CTTCGCAATG AAAAGAAACC GATCGGTTCT TTCCTTTTCC TCGGCCCCAC GGGGGTAGGC
                                                                              1800
20
      AAGACCTATT TGGCCAAGAA GCTCGCCGAA TACCTGTTCG AGGATGAGAA TGCCATGATC
                                                                              1860
      AGGGTGGATA TGAGCGAGTA TATGGAGAAG TTCTCCGTTT CGCGTCTCGT GGGTGCCCCT
                                                                              1920
      CCGGGATATG TGGGCTATGA AGAAGGCGGC CAACTGACGG AGCGCGTAAG ACGCAAACCC
                                                                              1980
      TATTCCGTGG TTCTCTTGGA TGAGATCGAA AAGGCGCATG CCGATGTCTT CAATCTGCTC
      TTACAGGTGA TGGACGAAGG TCAGCTGACC GACAGTCTGG GACGGCGCGT GAATTTCAAG
                                                                              2100
25
      AACACCGTGA TCATCATCAC CTCCAACGTG GGTACACGCC AGCTCAAAGA CTTCGGGCAG
                                                                              2160
      GGTATCGGGT TCCGTICGGA AAAAGACGAG GAAGCGAACA AGGAGCATAG CCGTTCCGTG
                                                                              2220
      ATCCAAAAAG CTCTGAACAA GACGTTCAGC CCCGAATTTC TCAACCGTTT GGACGATATC
                                                                              2280
      ATCCTCTTCG ACCAACTGGG CAAGACGGAG ATTCGCCGGA TGGTGGACAT AGAGCTTAAA
                                                                              2340
      GCCGTCTTGG CGCGCATCCA TCGTGCCGGA TACGACCTCG TCCTTACCGA TGAAGCCAAG
                                                                             2400
30
      GATGTGATAG CGACGAAGGG ATACGACCTC CJATACGGAG CACGACCGCT CAAGCGCACA
                                                                              2460
      CTCCAGAACG AAGTGGAGGA TCGCCTCACG GATCTTATCC TCTCCGGACA GATCGAGAAA
                                                                             2520
      GGGCAGACGC TTACGCTCTC TGCTCGCGAT GGCGAGATCA TCGTACAAGA ACAAGCA
                                                                              2577
35
      (2) INFORMATION FOR SEQ ID NO:165
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1251 base pairs
                 (B) TYPE: nucleic acid
40
                 (C) STRAHDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
45
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
50
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...1251
55
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:165
      ATGAATTACC TGTACATACT GATTACACTT TTACTCTCCG GCTTTTTTTC CGGTGCTGAG
      ATTGCTTTCC TTTCTTCGGA CAAACTGCGT CTTGAGTTGG ACAGGAATAG AGGCGATCTC
                                                                               120
60
      ACAGGCAGAG CGTTAAACTT GCTGTATCGA CATCCGGACC AACTGGTGAC TACTCTCCTT
                                                                               180
      GTGGGTAATA ATATCGTTTT GGTAGTCTAT GGTCTGCTGA TGGCGGGATT GCTGGCCGCA
                                                                               240
      CCTTTGGCGC AATGGATTGA TAACGATGCT ATGATCGTCG TTCTCCAATC TGTCTTATCC
                                                                               300
      ACTATCATCA TACTGTTTAC CGGGGAATTT CTACCCAAAG CCATTTTCAA GACCAATGCC
                                                                               360
      AATATGATGA TGAGGGTATT CGCCCTCCCT ATCGTAGCGA TCTATTATCT GCTTTATCCT
                                                                               420
65
      CTGTCTAAAC TCTTCACCGG TTTATCTCGC TCTTTTATTC GTCTGGTGGA CAAGAATTAT
                                                                               480
      GTGCCTACAA CAGTAGGGTT GGGGCGCGTA GATCTCGATC ATTATTTGGC AGAAAATATG
                                                                               540
      TCCGGAGAAA ACGAACAGAA CGACTTGACT ACCGAAGTGA AAATCATCCA GAATGCGCTG
                                                                               600
      GATTTTTCCG GTATTCAGGT GCGAGACTGC ATGATCCCAC GCAATGAGAT GATAGCATGT
                                                                               660
      GAGTTGCAAA CGGATATTGA AGTACTCAAA ACGACTTTTA TCGATACCGG TTTGTCCAAG
70
      ATCATTATCT ACAGACAGAA CATAGATGAC GTAGTAGGAT ATATCCATTC GAGCGAAATG
                                                                               780
      TTTCGTGGGC AAGACTGGCA AAAACGTATC AATACTACTG TATTCGTACC CGAAAGCATG
                                                                               840
      TATGCCAATA AACTGATGCG ACTACTCATG CAGCGCAAGA AAAGCATTGC GATCGTCATC
                                                                               900
      GATGAACTIG GAGGTACGGC CGGAATGGTC ACATTAGAGG ATTIGGTAGA AGAGATITIC
                                                                               960
      GGTGACATTG AGGACGAACA CGACACTCGC AAGATCATAG CCAAACAGCT CGGCCCTCAT
                                                                             1020
75
      ACCTATCTGG TCAGTGGTCG TATGGAAATA GATGATGTGA ACGAACGTTT TGGGTTGTCC
                                                                             1080
```

WO 99/29870 PCT/AU98/01023

| | TTGCCTGAGT CTGACGACTA CCTTACCGTG GCCGGATTTA TCCTGAATAG CCATCAAAAT ATCCCACAGG CCAATGAGGT CGTGGAGATT GCTCCTTATA CTTTTACCAT TCTCAGATCT TCTTCCACCA AGATCGAACT GGTGAAAATG TCCATCGACG ACCAATCGAA C | 1140 1200 1251 |
|------------|---|---------------------------------|
| 5 | (2) INFORMATION FOR SEQ ID NO:166 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 879 base pairs (B) TYPE: nucleic acid (C) STRANDEDMESS: double (D) TOPOLOGY: circular | |
| 15 | (ii) HOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 25 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1879</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:166 | |
| 30 | ATGAAACAGA ACTACTTCAA AAGAGTCTGC TCACTGCTTT GGCTGGTTTT ACCCATGCTT ATTATGCCAT TGGAAGTAGC AGCTCAAGAG ATTATTCCGA ACGAAGAGT GTTGGAATCA TTGACTTTCG TTGCACCGGT TGAGGAGACA GACGCAATAG AGGCAGAGGT AGAAGCTCTG CAGGAGATAG TCGCTACTGA GGAGATTGCG GAGCAGCTG TTCGTTCTTA TACCTACACG GTCTATCGTG ATGGCGTGAA GATTGCTTCA GGATTGACTG AGCCCACTTT TCTCGATGAA | 60 120 180 240 300 |
| 35 | GATGTTCCTG CCGGCGAACA TACCTACTGC GTAGAAGTAC AGTATCAGGG AGGCGTATCC GACAAAGTAT GCGTGGACGT AGAGGTGAAG GACTTCAAAC CGGTTACCAA TCTCACCGGA ACTGCTTCCA ATGACGAAGT TTCTTTGGAC TGGGACGGTG TGGAAGAGAA AGCTGAAGAG CCGGCAAGTG ATAAAGCAGT CAGCTACAAC GTCTACAAGA ATGGAACCTT GATCGGTAAT ACAGCTGAAA CTCATTATGT GGAGACCGGT GTAGCCAATG GTACATACAT CTACGAAGTG | 360 420 480 540 600 |
| 40 | GAAGTAAAGT ATCCTGACGG TGTATCTCCG AAGGTGGCTG TAACCGTGAC CGTGACCAAC AGCTCATTGA GCAATGTAGA TGGACAGGCT CCTTACACAT TGCGAGTAGA AGGCAAGAAG ATTATTGCGG AAGCCCATGG TATGATCACG CTCTACGACA TCAACGGACG TACCGTGGCC GTAGCCCCGA ATCGATTGGA ATACATGGCG CAAACCGGTT TCTATGCAGT GCGCTTCGAT GTGGGGGAATA AACACCATGT ATCGAAAATA CAAGTAAGA | 660 720 780 840 879 |
| 45 | (2) INFORMATION FOR SEQ ID NO:167 | |
| 50 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 936 base pairs (B) TYFE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 55 | (ii) MOLECULE TYPE: DNA (genomic) | |
| 30 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 60 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| 65 | (ix) FEATURE: (A) MAME/KEY: misc_feature (B) LOCATION 1936 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:167 | |
| 70 | ATGATTCCTC TAAGCGAGAG TTTTGAATCA GGTATTCCAG CTATATGGAA GACCATTGAC GCAGATGGCG ATGGCTATAA TTGGATGCAT TCACGGGACA GAGTGGTCTC TGTGTCTCTT CGGCTTCATA CATAGGCGGC GTCGGAGCTT TGACTCCGGA CAATTATCTG ATAACACCCG AATTAAAACT ACCCACAGAC GCGTTGGTGG AAATAATCTA TTGGGTATGT ACTCAAGATC TCACTGCTCC ATCGGAGCAC TATGCCGTTT ATTCCTCTTC TACAGGCAAT | 60 120 180 240 300 |
| <i>7</i> 5 | AATGCTGCTG ACTTTGTTAA TCTCTTATAT GAAGAGACTT TGACTGCCAA ACGGATACAA TCCCCCGAGT TGATCCGCGG AAATCGGACA CAAGGTGTTT GGTATCAAAG AAAGGTGGTA | 360 420 |

PCT/AU98/01023

```
CTCCCTAACG ATACTAAATA TGTTGCTTTC CGCCATTTTA ATTCCACGGA TAATTTCTGG
      CTCAATTTGG ATGAAGTATC TATCCTGTAT ACCCCTCTTC CCCGAAGAGC TCCGTGTCCG
                                                                                540
      CATCCGGGTG GTTACACTTA TTCTGTATTC CGTGATGGAC AAAAGATAGC GAGTGGATTG
                                                                                600
      TCGGCATTGG CATATATCGA TACGGATGTA CCGTATGGGA CTCAAGACTA TTGTGTCCAA
                                                                                660
      GTCAATTATC TGCAAGGAGA CTCGTATAAA GTCTGCAAAA ATATAGTGGT GGCAAATTCT
                                                                                720
      GCAAACATCT ATGGGGCGGA TAAGCCTTTT GCGTTGACCG TGGTTGGCAA GACCATTGTA GCGAGTGCTT TCAAAGGAGA GATCACTCTT TATGACATTC GTGGCCGGCT GATAGCTTCC
                                                                                780
                                                                                840
      GGCTGCGATA CGCTTAGGTA CAAAGCGGAA AATGGTTTTT ACCTCATTAA AATACAGGTA
                                                                                900
      AACGGAACTG TCTATACTGA GAAAATCCAA ATCCAA
                                                                                936
10
      (2) INFORMATION FOR SEQ ID NO:168
            (i) SEQUENCE CHARACTERISTICS:
15
                 (A) LENGTH: 2529 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
20
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
25
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROHOMAS GINGIVALIS
           (ix) FEATURE:
30
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...2529
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:168
35
      ATGAAAAAA GTTTTCTTTT AGCCATAGTA ATGCTCTTTG GCATTGCCAT GCAGGGACAT
      TCTGCTCCGG TTACGAAAGA GCGAGCTTTG AGTCTGGCTC GGCTGGCTTT GCGACAGGTA
      TCCTTGCGAA TGGGACAAAC AGCAGTATCT GACAAGATTT CCATCGATTA CGTTTATCGG
      CAAGGAGATG CTGAGAGGGG TATCACATCA CAAGAGGAAG GCTCTCCTGC ATATTTTAT
                                                                                240
      GTAGCTAATC GTGGAAATAA TGAGGGCTAT GCTCTTGTAG CAGCAGATGA CAGAATACCG
                                                                                300
40
      ACAATTTTAG CCTATTCACC CATTGGCCGT TTCGACATGG ACAGTATGCC GGACAATCTT
                                                                                360
      CGCATGTGGC TACAAATTTA CGATCAGGAA ATAGGCCTGA TACTTTCCGG AAAAGCTCAG
                                                                                420
      CTCAATGAAG AGATATTACG TACCGAGGGC GTACCGGCTG AAGTACATGC TCTGATGGAT
                                                                                480
      AACGGTCATT TTGCCAACGA TCCCATGCGA TGGAATCAAG GTTACCCATG GAACAATAAG
                                                                                540
      GAACCACTGC TTCCTAATGG CAATCATGCC TATACCGGCT GTGTTGCTAC TGCTGCAGCA
                                                                                600
45
      CAAATCATGC GCTACCATAG CTGGCCGCTT CAAGGTGAAG GCTCTTTCGA TTATCATGCA
                                                                                660
      GGTTCATTAG TTGGCAACTG GTCCGGCACA TTTGGTGAAA TGTACGACTG GATCAATATG
                                                                                720
      CCCGGAAATC CCGACCTTGA TAATCTGACT CAATCTCAAG TGGATGCCTA CGCCACACTG
                                                                                780
      ATGCGTGATG TGAGTGCATC TGTTTCGATG AGTTTTTATG AAAATGGAAG TGGTACGTAC
                                                                                840
      AGCGTTTATG TAGTAGGAGC CTTGCGAAAC AACTTTCGCT ACAAGCGTTC ACTGCAGCTA
                                                                                900
      CATGTACGCG CCTTATATAC CTCACAGGAG TGGCACGATA TGATCCGCGG GGAACTTGCC
TCCGGAAGGC CGGTCTATTA TGCAGGGAAT AACCAGAGCA TAGGACATGC TTTCGTTTGC
50
                                                                                960
                                                                               1020
      GATGGTTATG CTTCGGATGG TACTTTCCAT TTCAACTGGG GTTGGGGAGG TGTTTCCAAC
                                                                               1080
      GGCTTCTACA AACTAACACT CCTCTCGCCG ACTTCGTTGG GTATCGGAGG TGAGGGAATA
                                                                               1140
      GGTTTTACCA TTTATCAAGA GATCATCACC GGTATCGAAC CGGCTAAGAC TCCCGCTGAA
                                                                               1200
55
      GCCGGTACAG ATGCCTTGCC GATCTTGGCA CTGAAAGACA TAGAAGCCGA GTATAAAAGT
                                                                               1260
      GAATCCGGAT TGAACGTAGG GTATTCGATA TATAATACAG GTGAAGAGCA ATCAAATCTT
      GACCTCGGAT ACAGATTGAA CAAGGCTGAC GGAGAAGTCA TAGAGGTGAA AACTTCATCT
      ATCAATATCT CTTGGTACGG ATACGGAGAG CATCCCGAGA GTTTCTCATT GGCACCTAAT
      CAGTTGTCAC AAGGAATCAA CACCATCACC CTACTTTATC GTCGCACAGG CACCGAACAG
60
      TGGGAGCCGG TACGGCATGC ACAGGGAGGA TATGTCAATA GCATTAAAGT AAATACGACA
      GACCCGAACA ATGTCGTAGT CACGGTAGAT AATAACGAAG GCAAGCTCAG TATCGTCCCC
                                                                               1620
      AACAGCTTTG TCGCAGATCT GAATTCTTAT GAACATAGTA CGATTACAGT ACAGTTCAAT
                                                                               1680
      AGCGACAGCC CTGATGAGAT CCGTACACCC GTAGCCTTTG CTCTATCTAC AGGAGCTACT
                                                                               1740
      GCGGACGATG TAATATCTTT GGGCTGGGTA ATGGCTGAAG TTCCGGGCGG TAGCAGCAAC
                                                                               1800
65
      TATCCGGTGG TTTGGTCTAA AGACGTTCTC ACTCTCTCGG AAGGCGACTA TACATTGTGG
                                                                               1860
      TATAGATTTT CCATCAACAA CCAAAAGGAT GAATGGAAAA AGATCGGAAG CGTGTCAGTA
                                                                               1920
      AAAACACCGA CAGAGTATAC GCACCCCTTA TTCGAAGTGG GCCATAATCA AACTTCTACC
                                                                               1980
      TATACGCTGG ATATGGCACA CAACAGAGTA TTGCCCGACT TTACACTCAA AAATCTCGGA
                                                                               2040
      TTGCCTTTCA ATGGTGAGTT GGTTGTTGTT TTCCGCCAAA CACAATCCTC ATCGGGGTCT
                                                                               2100
70
      TTATGGGCAG CTCAAGAAAC AGTACATATC AAGCAAGGAG AAACTTTCGT ATATAAACCT
                                                                               2160
      GTTGTCGAAG GCCCTATACC TGATGGATCC TATCGTGCGA CCCTCCATGC ATTCGTAAAC
                                                                               2220
      GGACAACAAC AGTTGTACCT CAAGGGGAAA AGGAACTACA CGGTGAAGAT CGTCAATGGT
                                                                               2280
      ACAGCGGTAG AAGCAATAGA ATCGTCAGAA GAGATCAGAG TATTCCCTAA TCCGGCACGC
                                                                               2340
      GATTATGTGG AAATATCGGC ACCTTGCATT CCCCAAGAAA CATCTATCAT TCTTTTCGAT
                                                                               2400
75
      CTGTCAGGCA AGATTGTCAT GAAGAATAGT TTATCAGCGG GGCATGGCAG AATGGATGTC
                                                                               2460
```

| | AGCCGACTTC CTAATGGGGC CTACATCCTT AAGGTGGATG GATATACGAC GAAAATAAAT ATAGTGCAC | 2520 2529 |
|------------|---|--|
| 5 | (2) INFORMATION FOR SEQ ID NO:169 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 870 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 15 | (iii) HYPOTHETICAL: MO | |
| | (iv) ANTI-SENSE: NO | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 25 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1870</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID HO:169 | |
| 30 | ATGAAAAAGC TATTTCTCTC GCTCACGAGT CTTGTAATGG TCTTCGCTGT TGCAAGTTGC GATATAATCG ACAAGGATCA AACCCTCTTG CCGGCTCCGA CCAATGTGAC ACCCGATAAT CCGGATGACA ATCCTTCGGA GATCGACATT ACGCAGACGC ACACAGAAAA ATATGTTTTG GCTGAAGGAAT TTACCGGCCA AAAATGTCTC AACTGTCCGA AAGGTCATCG CAAACTGGCG GCTCTCAAGG AGCAATACGG TAAGAGATTG ACTGTTGTCG GTATACATGC CGGCCCTGGA TCTCTCGTGC CACCTCTTTT CCGTACAGAA GCCGGAGACG CATATTATAG CAAGTTCGCC | 60 120 180 240 300 360 |
| 35 40 | AATAATACCC CTCTCCCTGC GCTGATGGTT TCGCGCAAAA AGTTCGGCTC TTCCTACGTT TATGATAAGA GCTACAAAAC GTGGGACGTG CCTATTGCCG AGCAGATGGA GCAAAAGGCG AAGATCAATA TCTTTGCCGT GGCCGAATAC ACCGATACCC AAAAAGATCAA GGTGACTGTA AAGGGTAAAA TACTGGAGGG GAATACACTC CCGAAGTCCA TGGTTCAGGT GTATCTGTTG GAGGATAAGC TGATCGCTCC GCAGGTGGAT GGCAATACGA CAGTCGAGAA TTACGAGGCAC AATCACGTGT TGCGTGGAGC CGTTAATGGT ATTTGGGGCG AAGAATTTGT GAATCTCAAA | 420 480 540 600 660 720 |
| 40 | GATTATTTGT ATACTTACGC CGTTGAACCG CTCTCGGGTA TGTCCTTCGT AGCCGAGAAT TATTCGATTG TGGCTTTTGT ATACGATGTG CAGACGTTCG AAGTGTATGA CGTTGTGCAT GTAAAGATCA ATCCGCAATC CGATGGCAAA | 780 840 870 |
| 45 | (2) INFORMATION FOR SEQ ID NO:170 | |
| 50 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 669 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 55 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 60 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROHONAS GINGIVALIS | |
| CE | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1669 | |
| 65 | (%i) SEQUENCE DESCRIPTION: SEQ ID NO:170 | |
| 70 | ATGAAGAAAT CAAGTGTAGT AGCCTCAGTT TTGGCCGTGG CTCTCGTGTT CGCCGGTTGC GGACTGAACA ATATGGCAAA AGGCGGCCTT ATCGGCGCCG GAGTAGGAGG TGCCATTGGT GCCGGAGTAG GTAACGTAGC CGGAAATACG GCTGTCGGTG CCATCGTCGG TACTGCAGTC GGTGGAGCAG CCGGTGCTCT CATCGGAAAG AAGATGGACA AGCAGAAAAA AGAACTGGAG GCCGCAGTAC CCGATGCTAC GATTCAGACA GTAAATGACG GAGAGGCTAT TCTGGTTACT | 60 120 180 240 300 |
| <i>7</i> 5 | TTCGATASCG GTATCCTCTT TGCGACGAAC TCCAGCACT TGAGTCCCAA CTCACGCACT GCGCTGACGA AGTTTGCTGC AAACATGAAC AAAAACCCCG ACACGGATAT TCGTATCGTA | 360 420 480 |

| 5 | GCCAGCGTAT ATTCTTTCCT GAATTCTCAG GGTGTGAGTA TGTCGCGCAT GGCAGCCGAA GGGCGTGGGA GCCATGAACC GGTTGCAGAC AATAGCACAG TTGCCGGACG TTCGGCCAAC CGCCGTGTGG AGGTTTATAT CTTGCCGAAT GCCAAGATGA TCGAACAAGC ACAGCAAGGT ACGCTGAAG | 540 600 660 669 | | | | |
|------------|--|---------------------------------|--|--|--|--|
| | (2) INFORMATION FOR SEQ ID NO:171 | | | | | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1011 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | |
| 1 5 | (ii) MOLECULE TYPE: DNA (genomic) | | | | | |
| | (iii) HYPOTHETICAL: NO | | | | | |
| 20 | (iv) ANTI-SENSE: HO | | | | | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | | | | | |
| 25 | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 11011</pre> | | | | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:171 | | | | | |
| 30 | ATGTCGAAAA AATCGATCCT TCTGCTTTGC TGTTCGCTGT GCTTCATTTC TGCTACGAAG GCTGTGACCC CCGTCAGAAA TGTGCGCAAT AGCCAAGTGA ACAGCAAAGC AAAGACCGAA CGTACAAAGC CCTCGGACTC TGTACGGTAC ATTAGCAACA TGATTGCAGA TCGCTGGAG TTCCGCAACA AGATTTCTTC CGAAAAAGAG GTAAGAAAAG CCGAATATGA AAATCGGCTG | 60 120 180 240 | | | | |
| 35 | GCGATGGAAG CACTCAATTA CCCTGCCATA GATTTATATG GTGAAGATTC TTGGAGCGAG TATGTAAACC CTTTCGTGGG TGCAGGAACC GATGTCGAAA TTCCGAACTC CTATGACATT GATTGCTCTT CGTTCGTGAT GCCCGTCGAA GATAAGCAGG TCACCTCTCA ATTTGGCTAC CGTCGGCGTT TCGGACGGAT GCACTATGGT ATTGATCTTT CAGTGAATCG TGGCGATACG ATACGAGCAG CCTTTGACGG GAAAGTTCGT GTACGCAGCT ATGAAGCGC TGGCTATGGC | 300 360 420 480 | | | | |
| 40 | TACTACATAG TCTTGCGCCA TCCGAACGGA CTGGAGACTG TGTACGGACA CATGAGTCGC CAATTGGTAG ACGAGAATCA GATCGTTCGA GCAGGACAAC CGATCGGATT AGGAGGCAGC ACGGGTCGAA GCACCGGTCC TCATCTTCAC TTCGAGACCC GCTTCATGGG TATTCCCATC AATCCGAGTA CCATTATAGA CTTCGATAAC GGAGTGCCGC TCCGAGACAT TTACACATTC | 540 600 660 720 780 | | | | |
| 45 | AAACGAGGA GCAATTCTCG CTATGCAAAA GCCTCTAAGA CTTCTTCTCG CTATGCAAAA AAAGGGAAGA AAGGCAGACA AGCTTCTTCT CCTATGACCT ATAGAATCAA AAAAGGCGAT ACTTTGGAAA CAATAGCCAA AAGGCACGGC ACTTCTGTTC AGAAACTCTG TGCTACCAAT GGCATTGGCA AGAGTAAAAT TTTGACTCCG GGCAAAGCCT TGAGGATCAA A | 840 900 960 1011 | | | | |
| 50 | (2) INFORMATION FOR SEQ ID NO:172 | | | | | |
| 55 | (i) SEQUENCE CHARACTERISTICS: (A) LEMGTH: 453 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | |
| | (ii) MOLECULE TYPE: DNA (genomic) | | | | | |
| • | (iii) HYPOTHETICAL: NO | | | | | |
| 60 | (iv) ANTI-SENSE: NO | | | | | |
| e e | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | | | | | |
| 65 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1453 | | | | | |
| 70 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:172 | | | | | |
| 75 | ATGGCAAAAA TCAATTTCTA TGCTGAAGGC GTCAGCCTTC CTCGGATCAG AAGACGGATC GTCGGTAAGT GGATAGCCGA AGTATGCAGC CGATATGGGA AAGCCGGTGG AGAAATCTCC TATCTTTCT GTGATGACG ATATATCCTG AAAGCCAATC AGGAATTTCT CGATCATGAC TACTACACCG ACATCATCAC CTTCGATTCC TGCGAAGCGG ATACGGTGAA TGGCGACCTG | 60 120 180 240 | | | | |

```
CTTATCAGTC TCGATACCGT ACGCTCGAAT GCCCGTGCTC TTGATCTTCG ATACGAAGAC
                                                                                  300
      GAACTGCATC GTGTCATTAT CCACGGCATA CTGCATCTTT GCGGATTGAA AGACAAGAGC
                                                                                  360
      AAAAAGGATG AAGCCCAAAT GCGTGCAGCC GAAGAGAAAG CCCTTGTCAT GCTGCGAGAA
                                                                                  420
      ACCATCGGAT CGGAGCTTTC CCTATTGCAT ACA
 5
      (2) INFORMATION FOR SEQ ID NO:173
            (i) SEQUENCE CHARACTERISTICS:
10
                 (A) LENGTH: 1173 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
15
          (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
          (iv) AUTI-SEUSE: HO
20
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
25
                 (A) HAME/KEY: misc feature
                 (B) LOCATION 1...1173
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:173
30
      ATGAAGGTAA AGTACTTAAT GCTCACATTG GTTGGAGCAA TTGCACTGAA CGCAAGTGCA
      CAGGAGAATA CTGTACCGGC AACGGGTCAG TTACCCGCTA AGAATGTTGC TTTTGCTCGC
                                                                                 120
      AATAAAGCAG GCAGCAATTG GTTTGTAACA CTGCAAGGCG GTGTTGCAGC GCAGTTCCTC
                                                                                 180
      AATGACAACA ACAACAAGA CCTCATGGAC CGCTTAGGAG CCATAGGTTC TCTTTCTGTC
                                                                                  240
      GGAAAGTATC ACAGCCCTTT CTTTGCAACT CGTTTGCAAA TTAACGGAGG TCAAGCCCAC
                                                                                  300
35
      ACTITCCTCG GAAAAAATGG CGAACAAGAA ATCAACACCA ATTITGGTGC AGCTCACTTC
                                                                                  360
      GACTTTATGT TTGATGTGGT TAACTACTTT GCACCATATC GCGAAAATCG TTTCTTCCAT
                                                                                  420
      TTAATTCCAT GGGTAGGTGT TGGCTACCAA CACAAATTCA TCGGTAGCGA ATGGAGCAAA
                                                                                  480
      GACAATGTGG AATCACTGAC GGCGAATGTA GGAGTTATGA TGGCTTTCAG ATTAGGAAAG
                                                                                  540
      CGAGTAGACT TTGTGATCGA AGCACAAGCA GCTCACTCCA ATCTCAATCT AAGTCGCGCA
                                                                                  600
40
      TACAATGCCA AGAAAACTCC CGTATTCGAA GATCCCGCAG GACGTTATTA CAATGGATTC
                                                                                  660
      CAGGGGATGG CTACAGCAGG TCTTAATTTC CGCCTGGGAG CCGTAGGCTT CAATGCCATT
      GALICCAATGG ACTACGCACT TATCAATGAT CTGAATGGTC AGATTAACCG TTTGCGCAGC
                                                                                  780
      GAGGTCGAAG AACTCTCAAA ACGTCCTGTA TCATGCCCCG AATGTCCTGA AGTAACTCCT
                                                                                  840
      GTTACTAAGA CAGAAAATAT ACTGACGGAA AAAGCTGTAC TGTTCCGTTT CGACAGCCAC
GTTGTGGACA AAGATCAATT GATCAACCTG TATGACGTAG CTCAGTTTGT AAAAGAAACT
                                                                                  900
45
                                                                                 960
      AACGAGCCGA TTACCGTTGT TGGTTATGCT GATCCTACGG GTAATACTCA ATACAACGAG
AAATTGTCTG AGCGTCGGGC TAAAGCCGTT GTTGATGTTC TGACAGGTAA ATATGGTGTG
                                                                                1020
                                                                                1080
      CCTTCCGAAT TAATCTCTGT AGAATGGAAG GGCGACTCTA CGCAACCGTT CAGCAAGAAA
                                                                                1140
      GCTTGGAATC GTGTTGTAAT CGTTCGCTCC AAG
                                                                                1173
50
      (2) INFORMATION FOR SEQ ID NO:174
            (i) SEQUENCE CHARACTERISTICS:
55
                 (A) LENGTH: 1155 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
60
          (ii) HOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
65
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORTPHYROMONAS GINGIVALIS
          (ix) FEATURE:
70
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1155
           (::i) SEQUENCE DESCRIPTION: SEQ ID NO:174
75
      ATGACATACA GAATTATGAA AGCTAAATCT TTATTATTAG CACTTGCGGG TCTCGCATGC
                                                                                  60
```

```
ACATTCAGTG CAACAGCCCA AGAAGCTACT ACACAGAACA AAGCAGGGAT GCACACCGCA
       TTCCAACGTG ATAAGGCCTC CGATCATTGG TTCATTGACA TTGCAGGTGG AGCAGGTATG
                                                                                    180
       GCTCTCTCGG GATGGAATAA TGATGTAGAC TTTGTAGATC GTCTAAGTAT CGTTCCTACT
                                                                                    240
       TTCGGTATCG GTAAATGGCA TGAGCCTTAT TTCGGTACTC GTCTCCAATT CACAGGATTC
 5
       GACATCTATG GATTCCCGCA AGGGAGCAAG GAGCGTAACC ACAATTACTT TGGAAACGCC
                                                                                    360
       CACCTTGACT TCATGTTCGA TCTGACGAAC TATTTCGGTG TATACCGTCC CAATCGTGTC
                                                                                    420
       TTCCATATCA TCCCATGGGC AGGTATAGGA TTTGGTTATA AATTCCATAG CGAAAACGCC
                                                                                    480
       AATGGTGAAA AAGTAGGAAG TAAAGATGAT ATGACCGGAA CAGTTAATGT CGGTTTGATG
                                                                                    540
       CTGAAATTCC GCCTATCAAG AGTCGTAGAC TTCAATATTG AAGGACAAGC TTTTGCCGGA
                                                                                    600
10
       AAGATGAACT TTATCGGGAC AAAGAGAGGA AAAGCAGACT TCCCTGTAAT GGCTACAGCA
                                                                                    660
      GGTCTAACGT TCAACCTTGG CAAGACAGAG TGGACAGAAA TTGTTCCTAT GGACTATGCT
TTGGTCAATG ACCTGAACAA CCAAATCAAC TCACTTCGCG GTCAAGTGGA AGAGTTGAGC
                                                                                    720
                                                                                    780
       CGTCGTCCTG TTTCATGCCC TGAATGCCCT GAGCCTACAC AGCCTACAGT TACTCGTGTA
                                                                                    840
       GTCGTTGACA ATGTGGTTTA CTTCCGTATC AATAGTGCAA AGATTGATCG TAATCAAGAA
                                                                                    900
      ATCAATGTTT ACAATACAGC TGAATATGCG AAGACCAACA ACGCACCGAT CAAGGTAGTA
GGTTACGCTG ACGAAAAAAC CGGTACTGCG GCCTATAACA TGAAGCTTTC AGAGCGTCGT
15
                                                                                    960
                                                                                   1020
       GCAAAAGCGG TAGCCAAGAT GCTTGAAAAG TATGGTGTTT CTGCGGATCG CATTACAATT
                                                                                  1080
      GAATGGAAGG GCTCATCAGA GCAAATCTAT GAAGAGAACG CTTGGAATCG TATTGTAGTA
                                                                                   1140
      ATGACTGCAG CGGAA
                                                                                   1155
20
       (2) INFORMATION FOR SEQ ID NO:175
            (i) SEQUENCE CHARACTERISTICS:
25
                  (A) LENGTH: 570 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
30
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
35
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
40
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...570
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:175
45
      ATGGAATTTT TCATGTTATT CATAGCGGCG GTTTTCGTTA ATAACGTCGT GCTGTCGCAG
                                                                                     60
       TTCCTCGGTA TATGCCCATT CTTAGGCGTA TCGAAGAAGG TAGACACCTC AATCGGTATG
                                                                                    120
      GGTGCAGCCG TGACATTCGT ATTGGCACTG GCTACCTTGG TTACCTTCCT GATTCAGAAG
                                                                                    180
      TTCGTTTTGG ATCGTTTCGG ATTGGGCTTT ATGCAGACCA TTGCATTTAT TTTGGTCATT
                                                                                    240
      GCCGCCTTGG TGCAGATGGT GGAGATCATA CTCAAGAAAG TATCTCCTCC CCTCTATCAG
GCACTGGGTG TATTCTTGCC CTTGATTACG ACGAACTGCT GTGTGCTCGG TGTGGCTATT
                                                                                    300
50
                                                                                    360
       TTGGTTATCC AGAAGGATTA TACCCTGCTC CAGAGCTTCG TCTATGCAAT ATCCACGGCT
                                                                                    420
       ATCGGTTTCA CCTTGGCAAT GGTTACTTTC GCAGGTATTC GAGAGCAACT CGATATGACC
                                                                                    480
      AATCTCCCCA AAGCTATGAA GGGAATACCT TCGGCACTCT TGGCTGCCGG TATATTGGCT
                                                                                    540
      ATGGCTTTCA TGGGCTTCAG CGGTATCGCC
                                                                                    570
55
       (2) INFORMATION FOR SEQ ID NO:176
            (i) SEQUENCE CHARACTERISTICS:
60
                  (A) LENGTH: 558 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
65
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SENSE: NO
70
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
75
                  (A) HAME/KEY: misc_feature
```

(B) LOCATION 1...558

| (xi) SEQUENCE DESCRIPTION: | SEQ | ΙD | NO:17 | 6 |
|----------------------------|-----|----|-------|---|
|----------------------------|-----|----|-------|---|

| 5 | ATGTTATTCA | TAGCGGCGGT | TTTCGTTAAT | AACGTCGTGC | TGTCGCAGTT | CCTCGGTATA | 60 |
|----|------------|------------|------------|------------|------------|------------|-----|
| | TGCCCATTCT | TAGGCGTATC | GAAGAAGGTA | GACACCTCAA | TCGGTATGGG | TGCAGCCGTG | 120 |
| | ACATTCGTAT | TGGCACTGGC | TACCTTGGTT | ACCTTCCTGA | TTCAGAAGTT | CGTTTTGGAT | 180 |
| | CGTTTCGGAT | TGGGCTTTAT | GCAGACCATT | GCATTTATTT | TGGTCATTGC | CGCCTTGGTG | 240 |
| | CAGATGGTGG | AGATCATACT | CAAGAAAGTA | TCTCCTCCCC | TCTATCAGGC | ACTGGGTGTA | 300 |
| 10 | TTCTTGCCCT | TGATTACGAC | GAACTGCTGT | GTGCTCGGTG | TGGCTATTTT | GGTTATCCAG | 360 |
| | AAGGATTATA | CCCTGCTCCA | GAGCTTCGTC | TATGCAATAT | CCACGGCTAT | CGGTTTCACC | 420 |
| | TTGGCAATGG | TTACTTTCGC | AGGTATTCGA | GAGCAACTCG | ATATGACCAA | TCTCCCCAAA | 480 |
| | GCTATGAAGG | GAATACCTTC | GGCACTCTTG | GCTGCCGGTA | TATTGGCTAT | GGCTTTCATG | 540 |
| | GGCTTCAGCG | GTATCGCC | | | | | 558 |
| 15 | | | | | | | |

(2) INFORMATION FOR SEQ ID NO:177

| | (i) SEQUENCE CHARACTERISTICS: |
|----|-------------------------------|
| 20 | (A) LEMGTH: 2499 base pairs |
| | (B) TYPE: nucleic acid |
| | (C) STRANDEDNESS: double |
| | (D) TOPOLOGY: circular |

- 25 (ii) NOLECULE TYPE: DNA (genomic)
 - (iii) HYPOTHETICAL: NO
- (iv) AHTI-SENSE: HO 30
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
- (ix) FEATURE: 35
 - (A) NAME/KEY: misc_feature (B) LOCATION 1...2499

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177

| 4 0 | ATGAAACAAC | TAAACATTAT | CAGCTTCATC | ATTGCTTTCC | TATTCTTAGG | AACGAGCGCA | 60 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | TCGGCTCAGC | AATCGGGCGG | ATCCGTTACA | GGTACCGTAG | TGGACAAAAG | CTCAAAAGAA | 120 |
| | CCTATCGCAT | ACGTACAAGT | ATTCGTCAAA | GGAACCACTC | TCGGAACTTC | CACGGATGCA | 180 |
| | AACGGAAACT | ACTCGATCAA | GGGAATCCCT | TCGGGTAATC | AAACTATCGT | AGCCCGACTC | 240 |
| | ATGGGTTACT | CCACTTGCGA | AGAAAAAGTA | CATATAGAAA | AGGGTGGTTC | CCGCCACGTA | 300 |
| 4 5 | GACCTCTATC | TGACCGAAGA | GATTCTCTCT | CTCGATGGGG | TAGTGGTATC | TGCCAATAGA | 360 |
| | AACGAGACTT | | AGCACCCTCG | | TACTGTCGCC | GGAACTTTTC | 420 |
| | CTCAAAACCA | ACTCTACCAA | | GGACTTAAGT | TCCAGCCCGG | TCTGCGCGTG | 480 |
| | GAGGACAACT | | CGGTTTCAAC | CAAGTTCGTA | | CGAAGGAGCC | 540 |
| - 0 | TATTCGCAAA | | CAGCCATCCC | | | | 600 |
| 50 | TTGGAGCAGA | | TATGATCGAA | CGTGTAGAAG | TAATTCGCGG | TGGAGGTTCG | 660 |
| | | | | | TAATTACGAA | | 720 |
| | | CCGAGATCAG | | ATGACCTTCG | ACCACGCGAA | AGGGTGGGGG | 780 |
| | | ATACGACCCA | | | CGGAAGACCG | | 840 |
| ~ - | GTCATGGTAT | | CAACTACCGT | | | | 900 |
| 55 | | CCAATCTGCG | | | | TAAGACCGGT | 960 |
| | | AAGCAACCCT | | | | TGGTGGCGAC | 1020 |
| | | ATCCTCCTTT | | | | CTATATCAAT | 1080 |
| | | TCAAATTCGA | | AGCGGTGGCA | AGGATTTCTT | CAGTCTGTAT | 1140 |
| 00 | | AAGACGTTCA | | | GTGGCGACTA | | 1200 |
| 60 | | GAGCAGTTCA | | | ACGAATACAA | CGATGCTTTC | 1260 |
| | ACGGCTCTTA | | GACTACCAAG | GGATTCGATT | TGCAAGGAGG | AGGTATGTAC | 1320 |
| | CGTCATACCT | TCGGAGAAAA | | ACCGGCGGAC | | CTACGGCCAA | 1380 |
| | | GAAGCGGCTA | | | AGAATACCTC | TACTTTTAGT | 1440 |
| CF | | AGCTCGAATA | | | CCCTTATCGG | | 1500 |
| 65 | GACTATGTTC | | GGATGGCAAA | | | | 1560 |
| | | ACGTACGATA | | | | | 1620 |
| | | GCGCTCCTCA | | | | | 1680 |
| | ACTCCTATCA | | | | AACGTTCACG | AAGCATCAGT | 1740 |
| 70 | GCTTCTTTCG | | CAGAGCCGAC | | | GGGAGAAGCC | 1800 |
| 70 | TTCTCCACCT | TTATCAGCAA | TCAGTTCAAA | CCATCCGATA | AGGTCGAAAC | CACGAGCGAT | 1860 |
| | GGCAAAGAAT | GGATCATTCG | | | ATGGAGTATC | GAAGGTATAT | 1920 |
| | GGTGTGAATC | TGGAGGGAAG | | AACAAATCGT | TCGACCTCCA | GCTCGGCGGT | 1980 |
| | | GAAGCCGCTA | | TATACCGCTG | | _, | 2040 |
| 76 | | | | | | | 2100 |
| 75 | GTTGCTACGG | TACGTCCTAC | CGAGCACTTC | GCCATCAATC | TCTCCGGTAC | ATTCACGGGC | 2160 |
| | | | | | | | |

```
AAAATGGATG TAGTACACGA AGCCTATGAA GGCGATATTC CCGCAGAACA CATAGCTCCG
                                                                             2220
      GACGGATCGT TCGACTTTGA AATGAATGGT CAGCAATTCA AAGGTTTTGGC CGAAGGTCAT
                                                                             2280
      GCCAAGCTCG TCAAGACTCC GGCCTTCGCC GATATAGACC TCAAGCTGAG CCACGACTTC
                                                                             2340
      CACCTTGCTT CCACTATGAC CTTGGAATTG AATGCCGGAA TACAGAACAT ATTCAACAGC
                                                                             2400
 5
      TATCAGAAAG ACACGGACAA GGGACCGGGT AGAGCTTCTA CTTACGTATA CGGTCCTATG
                                                                             2460
      CAGCCCAGAA GGATTTTCGT CGGTACAAAG ATCAATTTC
                                                                             2499
      (2) INFORMATION FOR SEQ ID NO:178
10
            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 2673 base pairs
                 (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
15
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
20
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
25
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
                (B) LOCATION 1...2673
30
          (xi) SEQUENCE DESCRIPTION: SEO ID NO:178
      ATGTACAAAA AGATTATTGC CGTAGCAGCT CTCTTCTGCG CCAGCATAGG GATCCTGAAA
      GGACAGTCCT CGGATCTGAC CCCTCAGGAT ACTATATATA GCCCTGAAAT ATCCTATGCC
                                                                              120
      AAGCCTATTC ATAAGACCAT AGCATCTATT GAGATCGAGG GAATGAGGTC TTTCGATGAC
                                                                              180
35
      TTTGTCTTGC GCAATCTTTC AGGCTTGGCT GTAGGTGATG AAGTCCTGAT TCCTGGAGAT
                                                                              240
      GCCATGTCTG CTGCCGTGAA TAGAATTATG CGTCAGGGCT ACTTCTCAAA TGTGCGAATC
                                                                              300
      ATCGCGGATA AATATGTCGG CAATAAAGTC TATCTGAAAA TCATTGTCAC TGAACGTCCT
                                                                              360
      CGCATCAGTA AGGTTACTTT TAGCGGGGTA AAGAAGTCTG AGAGAAGA TCTTGAAATG
                                                                              420
      AAAATCGGTC TTCGCGAGGG GATTCAGATG ACCAGAAATA ATGAAGACAA GGTCAGGCAA
                                                                              480
40
      ATCGTACAGA AGTATTTTAG TGAGAAAGGT TATCGCGATG CCAGCATACG GATAACGCAG
                                                                              540
      GAACCGGATC TTTCCAAAGA TGGCTTTGTC AATGTGCTTA TCTCGATTGA GAAGAAAAGC
      AAAACCAAGG TGAATGAAAT TTATTTTCC GGCAACAAGG CCCTTAGCAA TCATAAGCTA
                                                                              660
      AGAATGGCGA TGAAGAACAC CAATGCCAAA TTCAGTCTTA GAAAGCATAT TCGCTCATCT
                                                                              720
      TTCTTGAAAC TTTTTAGTAC TCATAAGTTT GTGGAAGAGA GCTACCGTGA AGATTTGGTC
                                                                              780
45
      CGATTGATAG AGAAGTATCA GGAATATGGA TATCGTGATG CTGAAATACT GACCGACAGT
                                                                              840
      GTCGTGAAGG CTCCTGACGG CAAAAGAGTG GATATTTATC TCAACATCGA AGAGGGGCAG
                                                                              900
      AAGTATTATA TTAAGGATGT CAACTTTGTG GGCAATTCAC AATATCCATC GGAGTATTTG
                                                                              960
      GAACGAGTGC TCGGAATAAA ATCCGGAGAT GTGTACAATC AGAGACGATT GGCTAAGCGT
                                                                            1020
      CTCAATGAAG ATGAAGATGC TGTGGGGAAC CTGTACTATA ACAATGGCTA TATTTTTGCG
                                                                            1080
50
      TGGGTCGATC CCGTGGAAAC AAATGTAGTG GGGGATTCTG TTTCGCTTGA TATTCGTATA
                                                                            1140
      GCGGAGGGGA AGCAGGCCAA TATCAATAAG GTGATCATCA AAGGAAATAC TGTCGTGTAC
                                                                            1200
      GAAGACGTAG TACGCCGAGA GCTTTACACA AAGCCCGGCC AGCTCTTTAG TCGCGAGGAT
                                                                            1260
      ATCATTAACT CTATTCGTCT CATCAATCAG CTTGGGCATT TCGATGCCGA AAAATCTATT
                                                                            1320
      CCCCGTCCGA TTCCCAATCC CGAAACAGGA ACAGTGGATA TAGAGTATGA TTTGGTGCCG
                                                                            1380
55
      CGTAGCAGTG ACCAATTGGA GCTTTCTGTC GGTTGGAGTC AGTCCGGACT TCTGTTCCGA
                                                                            1440
      GGAGCCATTA AGTTCACGAA CTTCTCTGTC GGCAACTTGC TCCATCCCTC GATGTATAAG
                                                                            1500
      AAAGGGATCA TTCCGCAAGG GGATGGGCAA ACACTATCAC TGAGTGCTCA GACCAATGGA
                                                                            1560
      AAGTACTATC AGCAGTATAG TGTCACATTT ATGGATCCAT GGTTTGGGGG CAAGCGGCCG
                                                                            1620
      GATATGTTCA GCTTCAGTGC ATTCTATTCC AAGACTACGG CGATTGACTC CAAGTTCTAC
                                                                            1680
60
      AATAGCAATG CCGGCAACTA CTATAATGCC TACTATAATA GCTACTACAA CAACTATAAT
                                                                            1740
      AGTTATTACA ACGGTATGTC GAACTATACC GGCGACCTCT ATACTCAGGC CAGCGATCCG
                                                                            1800
      GATCGTTCGC TTCAGATGTT AGGTACTTCG ATCGGTTACG GTAAGCGTTT GACTTGGCCG
                                                                            1860
      GACAATTGGT TCCAGATTTA TACTTCTCTG AACTACACCT ACTATAGACT GCGAAATTGG
                                                                            1920
      AGCTACAATA CCTTCCAAAA TTTCCATCAT GGCTCGGCTA ATGATCTCAA CTTGGAGCTG
                                                                            1980
65
      CGTCTCTCTC GTACTTCCAT CGATAATCCT ATTTATACCA GAAGCGGATC GGATTTCATG
                                                                            2040
      GTTTCTGTTG CTGCTACTCT TCCTTATTCT TTGTGGGACA ATCATGACTA TGCCAGCCAG
                                                                            2100
      AACCTCAGCG TAAGCGATCG TTACAGATTT ATCGAGTATC ACAAGTGGAA GTTTAGAGGA
                                                                            2160
      CGAGTTTTTA CTCCATTGCT CAATCCTGCT ACGCATAAAT ATACACCGGT GCTCATGAGT
                                                                            2220
      CGAGTGGAAG GAGCAGTTCT TGGTTCGTAT AATTCCAATA AGAAATCTCC TTTCGGTACT
                                                                            2280
70
      TTCTATATGG GAGGTGATGG TATGTCCAGC TATTATGGTG GCTACATGAA TGAGACTATA
                                                                            2340
      GGTTTGCGTG GTTATAAGAA CGGATCTATT GCCGGTAATA ACTACGACTA TGCATATGCT
                                                                            2400
      TATATGCGGC TTACGATGGA ACTACGTTTC CCGATTCTGT TTGAAAACTC ATTCAATGCG
                                                                            2460
      TGGCTCTTAG CTTTTGCCGA AGCAGGCAAT GCGTGGCGCA GTATCGACAA TTATAATCCC
                                                                            2520
      TTTAACCTGA AGCGATCGGC CGGTGTAGGA TTGCGTGTAA CGTTACCGAT GGTCGGAATG
                                                                            2580
75
      CTCGGTATCG ATTGGGGATA TGGCTTTGAC CGTCCGGACA ATTCTCTACA GCGAGGAGGA
                                                                            2640
```

| | AGCAATGTCC ACTTTGTGCT CGGACAGGAG TTC | 2673 |
|-----|--|------------------|
| | (2) INFORMATION FOR SEQ ID NO:179 | |
| 5 | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 522 base pairs (B) TYPE: nucleic acid | |
| 10 | (C) STRAHDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DHA (genomic) | |
| 4.5 | (iii) HYPOTHETICAL: NO | |
| 15 | (iv) ANTI-SENSE: NO | |
| | (vi) ORIGINAL SOURCE: | |
| 20 | (A) ORGANISH: PORYPHYROHONAS GINGIVALIS | |
| | <pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1522</pre> | |
| 25 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:179 | |
| 20 | ATGAATGGCG ATATGAAACG GTTTTTGATT TTGATCGGCT TTGCACTGGC GGTAGCTTTC | 60 |
| | TCCGGTTTTT CCCAAAAGTT CGCTTTGGTA GATATGGAAT ATATCCTCAG GAATATTCCT GACTATGAGA TGATGAACGA ACAGCTGGAA CAGGTGTCCA AGAAATGGCA AAATGAAATC | 120 180 |
| 30 | GAAGCTCTCG AAAATGAAGC CCAATCTATG TATAAGAAGT ATCAGAGCGA TCTCGTATTC TTGTCTGCTG CACAGAAGAA AACCCAAGAA GAGGCTATCG TAAAGAAAGA GCAGCAAGCA | 240 300 |
| | TCCGAGCTCA AGCGGAAGTA TTTCGGCCCG GAGGGGGAGC TGTATAAGAA ACGCTCCGAT CTGATGAAGC CTATTCAGGA TGAGATTTGG AATGCTATCA AAGAGATTGC CAAGCGTAAC | 360 420 |
| 35 | AACTATCAGA TGGTGCTTGA TAGAGGTACG TCCGGAATTA TCTTTGCCAG TCCGTCTATT GACATTAGCG ACCTTGTACT GAGCAAGATG GGCTTTAGCA AG | 480 522 |
| | | |
| 40 | (2) INFORMATION FOR SEQ ID NO:186 | |
| 40 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 510 base pairs | |
| | (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 45 | (ii) NOLECULE TYPE: DNA (genomic) | |
| | (iii) HYPOTHETICAL: NO | |
| 50 | (iv) ANTI-SENSE: NO | |
| | (vi) ORIGIMAL SOURCE: | |
| FF | (A) ORGANISH: PORYPHYROHOHAS GINGIVALIS | |
| 55 | (ix) FEATURE: (A) NAME/KEY: misc_feature | |
| | (B) LOCATION 1510 | |
| 60 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:180 ATGAAACGGT TTTTGATTTT GATCGGCTTT GCACTGGCGG TAGCTTTCTC CGGTTTTTCC | 60 |
| | CAAAAGTTCG CTTTGGTAGA TATGGAATAT ATCCTCAGGA ATATTCCTGA CTATGAGATG ATGAACGAAC AGCTGGAACA GGTGTCCAAG AAATGGCAAA ATGAAATCGA AGCTCTCGAA | 60 120 180 |
| 65 | AATGAAGCCC AATCTATGTA TAAGAAGTAT CAGAGCGATC TCGTATTCTT GTCTGCTGCA CAGAAGAAAA CCCAAGAAGA GGCTATCGTA AAGAAAGAGC AGCAAGCATC CGAGCTCAAG | 240 300 |
| | CGGAAGTATT TCGGCCCGGA GGGGGAGCTG TATAAGAAAC GCTCCGATCT GATGAAGCCT ATTCAGGATG AGATTTGGAA TGCTATCAAA GAGATTGCCA AGCGTAACAA CTATCAGATG | 360 420 |
| | GTGCTTGATA GAGGTACGTC CGGAATTATC TTTGCCAGTC CGTCTATTGA CATTAGCGAC CTTGTACTGA GCAAGATGGG CTTTAGCAAG | 480 510 |
| 70 | | |
| | (2) INFORMATION FOR SEQ ID NO:181 | |
| 75 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 489 base pairs | |

```
(B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
                    (D) TOPOLOGY: circular
  5
             (ii) MOLECULE TYPE: DNA (genomic)
            (iii) HYPOTHETICAL: NO
             (iv) ANTI-SENSE: NO
 10
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
             (ix) FEATURE:
 15
                    (A) HAHE/KEY: misc feature
                    (B) LOCATION 1...489
             (xi) SEQUENCE DESCRIPTION: SEO ID NO:181
 20
        ATGAAGAAAT TTTTTCTCAT GCTTCTGATG GCTCTTCCTT TGAGCCTCTT GGCACAAAAG
        GTGGCAGTGG TAAACACTGA GGAGATCATT TCCAAAATGC CGGAACAAGT AGCTGCTACC
                                                                                         120
        AAACAGCTCA ACGAATTGGC CGAAAAGTAT CGCCTTGATC TCAAGAGTAT GGACGATGAG
                                                                                         180
        TTTGCCAAAA AGACAGAAGA ATTTGTAAAG GAAAAAGACT CTCTACTGGA GAACATCCGC
                                                                                         240
        AATCGTCGTC AGCAGGAACT TCAGGATATT CAAACTCGTT ATCAGCAGTC ATACCAAACG
                                                                                         300
 25
        ATGCAGGAGG ATTTGCAAAA GCGCCAACAA CAGCTTTTTG CTCCTATCCA ACAAAAGGTG
                                                                                         360
        GCTGATGCCA TCAAGAAAGT GGGTGACGAA GAAAACTGTG CCTACATCAT GGAGGCCGGT
                                                                                         420
        ATGATGCTTT ACACCGGAGC TACTGCTATT GACTTGACCG CAAAGGTAAA AGCGAAACTC
                                                                                         480
        GGAATCAAG
                                                                                         489
30
        (2) INFORMATION FOR SEQ ID NO:182
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 2481 base pairs
35
                   (B) TYPE: nucleic acid
                   (C) STRAHDEDNESS: double
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
40
            (iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
45
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
                   (A) HAME/KEY: misc_feature
50
                   (B) LOCATION 1...2481
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:182
       ATGAAGGAAG CTATTCCCCG AAAGAACAAG TATATAAAGC TCAACGGTAT ATACAGATTG TCATTCATTC TGCTATGCTG CCTGCTATGC TCTCAGGCAG CTATGGCACA AGGCGTCAGG
                                                                                         60
55
                                                                                        120
       GTATCGGGCT ATGTGCTCGA CCGTGGGGAA AAGCCGATCC CGTTCGCCGG AGTCAAAGTG
                                                                                        180
       CGTGGTACGG GGACAGGCGC AACGACGAAT CTGAAAGGAT ACTACGAGTT TCGGATGAAG
                                                                                        240
       GCCACGACGG ACAGCATCAC GATCGAGTTC AGCTCCATGG GGTACCAAGG GGTAAGTCGC
                                                                                        300
       AGCTTTCCGT CTCTGACCAA GGACACTCGG CTGAATGTTC GTTTGGCAGA GGCCGAGATG
                                                                                        360
60
       GAGCTTTCGA GCGTGACGGT ACAGGCCACA AAACGCAGAC TCAACACGAT GGAGCGCGTC
                                                                                        420
       AATACCCGAG ACCTTCGTGT CAATGCAGGG CCTACGGGAG GGGTGGAATC GCTCATCAGT
                                                                                        480
       ACCTACGCAG GAGTAACGCA GAACAATGAA CTAAGCTCGC AATACTCGGT TCGCGGAGGA
                                                                                        540
       AGCTACGATG AGAATATGGT CTATGTAAAC GGAGTGGAGG TTTATCGCCC GCTGCTGGTT
                                                                                        600
       CGCTCTGCAC AGCAGGAAGG TCTGAGCTTC GTCAATCCGG ATCTGACACA ATCCGTACAG
TTCTCCGCCG GAGGGTTCAC GGCCGACTAT GGCGACAAGA TGTCCTCCGT ACTGGATATT
CGCTACAAGC AACCGCAGGA GAAGGAAGGA GCGGTACTCC TCGGGATGCT ACAATCGAGT
                                                                                        660
65
                                                                                        720
                                                                                        780
       GCCTACTATG GCAGCAGTGC CGGAGCCTTC AGCCAAATCA CGGGTGTACG CTACAAGAGT
GCCAAATCGC TCTTGGGCAC TACGGACACG AAAGCCGAAT ACGATCCGAT CTATGCGGAC
GGACAGACAT TCATGACGTA CCGTTTCAGC CCCAAGCTGT CGGTTAGTTT CCTCGGCAAT
                                                                                        840
                                                                                        900
                                                                                        960
70
       ATTTCGCAAA CTCGCTACAA GTTTGTCCCT CAGACCCGTG AGACGAGCTT CGGTACACTG
                                                                                       1020
       AGCGATGCCA AAAAGTTGAA GATCTTTTTC GACGGTCAGG AACAAGATCG TTTCCTGACC
                                                                                       1080
       TACTTCGGTG CCTTCAGCAT GAACTTCGTG CCGGACGACA AACAGCGGCA TACGGTTACG
                                                                                       1140
       CTTTCGGCCT TCAACAGTAA CGAACGGGAG ACCTACGATA TTCAGGGAGA ATACTTTCTG
                                                                                      1200
       AACGATGTGC AGCTGGGGGC GGACGGAACT GCTTCGATGG CTTCGGGCTC AGAGAACTCC
                                                                                      1260
75
       AACGGCTTGG GCATCGGGCG CAATCACGAG CATGCGCGCA ACAGGCTGAG CTACCGCGTG
                                                                                      1320
```

```
CTGAACATGG GTTACAGAGG GGAGATGAAG CTGAACGAGA AGCATCGCCT GCAAGCCGGC
                                                                                      1380
        GTATCGGCAC AGATGGAGAA AATAGCCGAC CATATCAGCG AATGGGAACG GAGGGATTCG
                                                                                      1440
        GTAGGATACA ACCTACCTCA CTCGGAGACC GTATTGCTGA TGTACAATAA CCTATATGCC
        GATACGCAGA TGAGGGGAAC GCGCTTGTCG GCATTCGTAC AGGATCGATT CAACTTCAGC
                                                                                      1560
  5
       ATGGGAGGAG GTACATTTC TCTCATTCCG GGTATCAGAG CTTCGTGGTG GAGCTTCAAC
                                                                                      1620
        AAGGAGTTGC TCGTCAGCCC ACGTATCAGC GTGGGTTATT CTCCCGAAAG CAACCCGGCT
                                                                                      1680
        TTGGTACTGC GTGCAGCCGC CGGACTTTAT TATCAGGCAC CGTTTTACAA AGAGCTAAGG
                                                                                      1740
        CAGACGCATA AGGATGCCGA AGGCAATAAC GTGGTTGTCC TCAACGAGAA GATCCGCTCT
                                                                                      1800
        CAGGGAGCTT TTCACATTCT CGCAGGAGCA GACTATACCT TCGAAATGGG GGGGCGAAAA
                                                                                      1860
10
       TACAAGTTTA CGGCAGAGGC TTACTACAAG AGCCTGTTCA ACATCAACCC GTATATAATA
                                                                                      1920
       GAGAACGTGA AGATCCGCTA TCTGGGCGAA AACATCGGTT CGGGTTATGC TGCGGGTATC
                                                                                      1980
       GATCTCAAGC TCTTCGGCGA ACTGGTACCC GGAGTGGATT CGTGGCTGAC GGCTTCCATT
                                                                                      2040
       ATAAAAGCCC GTCAGAAACT GGATGGCTAC GGTTCTTTAC CACTGATGAA CGCACCCACT
TACAATTTCT CCTTCTTCCT TCAGGAGTAC GTGCCGGGCA ATAAACGCAT CACAGCCACC
                                                                                      2100
                                                                                      2160
15
       CTGCGGGCTG CACTAAGCGG AGGATTGCCC CAGCTCAATC CGAGCAAAGG GCTTAGCTCG CCGGCCTTTA CCGCACCGGC CTATAAGCGT GTCGATCTGG GGGTAATGTA CAAATGGCTC
                                                                                      2220
                                                                                      2280
       GACCCGGATG ACTCCTTTGC CGGCCGAAGC AAATGGCTAA TGGGAGTAAA AGGGGCCTAC
                                                                                      2340
       ATAGGGGCTG ACCTCTTCAA TCTGTTCGAC ATGACCAACG TCAATTCTTA CTACTGGGTG
                                                                                      2400
       TEGGATGEET ACCAACAGEA ATACGEEGTA CEGAACTACE TGACACGEEG CEAATTEAAC
                                                                                      2460
20
       CTGCGTCTCC TCGTCGAATT C
                                                                                      2481
       (2) INFORMATION FOR SEO ID NO:183
25
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2016 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRAHDEDNESS: double
                  (D) TOPOLOGY: circular
30
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
35
            (iv) AHTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
40
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...2\overline{0}16
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:183
45
       ATGTACAGCG GACATCATAA AATCCATTAT CCTTTTCTTA TCCTGTTGGT ATGCCTTGCT
                                                                                        60
       TTTGCTGCCT GCAAGAGCGT GAAGTTGAAA GATGCGGAGA AGGCACATGA TCGCCAAGAG
                                                                                      120
       TATACCAAGG CTGCCGATAT GTACAATACA TTATACAGGC GTACCCGACG AAAGCAGGTG
                                                                                      180
       GAGATGAAAG CTTATACGGC TTTCCGATCC GGTGAAAACT ATCGTGCCGC CGGCAGACAA
                                                                                      240
50
      GCCAAAGCTT TGCGTGGCTA TCTGAATGCC AGACGCTACG GGTATCCGGA TTCTGTGGTAC
CTGCTCCGTT TGGCACAGAC TTATCAGCAA GGAGGTAACT ATAAGGAAGC CGAGGTACTC
TTCCGTGGAT ATCTGGAAGC TTATCCGAAA AGTTATTTTG CAGCTATCGG TTTGGAGGGG
                                                                                      300
                                                                                      360
                                                                                      420
       TGTCTCTTTG CCCGCCAGCA AAAGGAATAT CCTACACGTT ACCGGATACG GCGAGCTGCC
       GASTGGAATT CGGCACGGGG CGACTTCGGC CCGGCCTATG CACCCGATGC TTCGGCTCTC
                                                                                      540
55
       TATTTCACAT CGAGCAGAAG CAAAGACGAC GGTTTGGATA ATAGCAGCAT AACGGGACTG
                                                                                      600
       AAACCCAACG ACATTTATAT CATCAAACGA GATGCACAAG GACGATGGGG ACGTCCCGAT
                                                                                      660
       AGCGTGTCCG GAGGAATCAA CACTCCATGG GATGAAGGCG TGCCAACGAT CACGCCCGAT
                                                                                      720
       GGTAGTACCA TATATTATAC GTTGGCGCAG CAAGGAGCCG ATTACGACCG TACGGTACAG
                                                                                      780
       ATCTATTCCG CCGCTCGGAG CGGAGAAGGC GGTTGGAGCA ACGGTTCGCT CGTGGACATT
                                                                                      840
60
       ATGCGCGATT CGCTCCGTAT GGCTGCTCAT CCCTCTATGT CGGCATCCGG CGATTACCTG
                                                                                      900
       TATTTCGTCA GCAATATAGG CGGTAGCTAT GGCGGCAAGG ATATTTATCG TGTCAAGGTG TCGGATCGTT CTTATGGTTC ACCGGAGAAT TTGGGGCCTG ATATCAATAC GCCGGGGGAC
                                                                                      960
                                                                                     1020
      GAAATGTTTC CCTTCATAGA TGGGGATAGT ACCCTTTTCT TCGCTTCGGA CGGACACGCC
GGTCTGGGAG GACTGGATAT TTTCAAAGCC ACGCTGGACT CTACCGGCCA ATGGCATGTA
                                                                                     1140
65
       GTCAATATGG GACAACCGGT CAATTCCTCT GCCGATGATT TCGGCTTGGC TGTGGAGCCT
                                                                                     1200
       AAAGGCAAAA ACAAAGAAGA AGCTTTGCCG GACAACGGAG TCAAAGGTGT ATTTTGTTCC
                                                                                     1260
       AACCGAGGCG ATGCACGCGG ATGGCCGCAC CTCTTCCATT TCGAACTGCC GGCTATCTAC
                                                                                     1320
       ACCGAGATTC AAGGTTATGT GATGGACAGA GAAGAAAATC CCATAGCCGG AGCCACTGTC
                                                                                     1380
       AGGATCGTAG GCGAACGCGG CCCCGTAGGA CAGGGATTCG TGACTACTCG TGACGATGGC
                                                                                     1440
70
       TCCTATAAGA TGAGCGTGCA GGGCGATACT CGCTATGTAA TGCTTGCCGG AGCATCGGGT
                                                                                     1500
       TATTTGAATC AGTACGTAGA ACTCAAGACC GATACCGCCA AGCAGAGTGA GACCTACTAT
                                                                                     1560
       GTGGACTTTT TCCTTGCATC GCGTGAGAAA GCCGAGGGCT TGCAAAATAT TTTCTATGAT
                                                                                     1620
       TTCGATAAAG CTACTCTTCG CCCCGAAAGC ATGAAGAGCT TGGACGAACT GATTCGTATC
                                                                                     1680
       CTCACGGACA ATCCGGATAT TCGGATCGAA TTGGGTTCGC ATGCCGACAG GAAAGGCCCC
                                                                                     1740
75
       GATGCTTACA ACCTCGGACT ATCTGACCGC AGAGCCAAAT CCGTGGTGGA TTACCTCACG
                                                                                     1800
```

```
AGTCGTGGCA TAGCGGCCGA CAGGCTTACG TGGAAAGGCT ACGGTAAGTC TGTCCCCAAG
                                                                                                1860
         ACGGTGACAG CCAAAATTGC CGAACGGCAC GATTTCCTGA AGGAAGGCGA TGTGCTCACC
                                                                                                1920
         GAGGAATTCG TAGCACCTTT GACCGAGGAG CAGCAGTCAG TCTGCGACCA ACTGAACCGT
                                                                                                1980
         CGTACCGAGT TCCGTGTGAT CGAAGAAGAG TTGCGT
                                                                                                2016
  5
         (2) INFORMATION FOR SEQ ID NO:184
               (i) SEQUENCE CHARACTERISTICS:
10
                     (A) LENGTH: 2124 base pairs
                     (B) TYPE: nucleic acid
                     (C) STRANDEDNESS: double
                     (D) TOPOLOGY: circular
15
             (ii) MOLECULE TYPE: DNA (genomic)
            (iii) HYPOTHETICAL: NO
             (iv) ANTI-SENSE: NO
20
             (vi) ORIGINAL SOURCE:
                     (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
             (ix) FEATURE:
25
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...2124
             (xi) SEQUENCE DESCRIPTION: SEO ID NO:184
30
        ATGAAAAAGT TTTTCTTCGC GCTACTATCG ATTGGTATTT CAGCGCAGGC TTTTGCCAAG ACGGACAACG TCCCGACAGA TTCGCTACGA GTACACAATC TTCAGACCGT CACGGTCTAT
                                                                                                  60
                                                                                                 120
        TCTACACGCA CGGCCGTACC TCTGAAAAAG ATACCGGCCA AGATGGAACT CATCTCATCG
                                                                                                 180
        CGCAACATCA AGCAGTCCGG CTTTAACAAC ATGACCGACA TCCTCAAGAC GCAAAGTTCG
                                                                                                 240
        CTCGATGTCA TACAATACCC GGGCTTTAGT TCGAACATCG GTATCCGCGG TTTCAAGCCC TCCGGCAAGT ATGTAACCGT ATTGGTAAAC GGCATCCCTG CGGGAACGGA CAATATCTCT
                                                                                                 300
35
                                                                                                 360
        ACGCTCAACA CGAGCAACAT CGAACAAATC GAGATCCTCA AAGGCCCGTT CTCTTCCATC
                                                                                                 420
        TACGGCACCA ATGCCATGGG CGGTGTGGTG AACATCATCA CCCACAAATC CAAGGACAAG
                                                                                                 480
        ATCCATGGCA ACGTTTCTCT CTTCGGCGGT AGCTACCAGA CCATGGCCGG ATCATTCAAC TTGGGTGGCC GCTTCGAGGA TATTTTCTCA TTCGATCTTA GTCTGGGCTT GGACAAGCAG
                                                                                                 540
                                                                                                 600
40
        AACAAGGACT ATAAGACCGG ATCAAACAAT TTCCTATCCC TGAGCAAACT GGAAGAAGCT
                                                                                                 660
        ATAGTAGATG TAAATGCTAC CAAAAACAAG AAAATGAAGG GGAGCGACTA TACTGTAGCA
                                                                                                 720
        ACGGGACGTC TGCGTTTCGG TATCGACTTC ACGCCCGAAT GGTCGCTGAA TCTGTATCAA AACGTATTCC TCGGAGATGC GATCCCCGTA GGAGGATCTA TATGGGGCGT TTACGGAGAA TCCCAAAAAAAA ATCTGAATCG TTCTTCGACC TCTTTCGAGC TGCTCGGCAA ACATGGCTGC
                                                                                                 780
                                                                                                 840
                                                                                                 900
45
        CACACGCTTC AATTCTCCCC CTACTTCAAC ATAGAGAAAT CGGAGAACTA TAACAATGCC GATCCCACCG GTTTCATCAA CTACAAAAGC GACTACTACA CCTATGGTGC CCTACTCCAG
                                                                                                 960
                                                                                               1020
        GACAAGATTT CCTTTGGAGG ACAAAATATC GTACTCGGTG TCGACAGCCG AAACATGACG
ATGGAGTCAG AAAGATTCGA GCAGGCAGGA GTGAATACAA AGCCATACAA CCCCGGATAT
                                                                                               1080
                                                                                               1140
        GCCACGAACA ATATCGGTTT GTTCGGACAG GCCAATTTCT ACCTGCTGAA CGATGCTCTA
                                                                                               1200
50
        TCGATATCTG CCGGTGCACG TGCCGACTTC ATGTTCTTTG ACCTGAAAGC GAACGAGTAT
                                                                                               1260
        CTCAACAATG AAGCCAAACA GGAAACTCAT AACGTAATCA ATCCGAATGT CGGAATCAAA
                                                                                               1320
        TATGAGTTTG TGAAAGGCCT TACAGCTCAT GGTACATTCG GTAGTGCATT CAGTGCTCCC
                                                                                               1380
        GATGCTTTCC AAAAAGCAGG CCAATACGTA GGCCCGTTCG GCACGACCAT AGGCAATCCT
                                                                                               1440
        GACCTGAAAC CCGAAAAGTC CATGACCTGG GACTTCGGTA TCGGATACAG CAATGCACGC
                                                                                               1500
55
       TGCGGGATCC AAGCCGACGT AACCTTAACC TATTTCCACA CCGACCACAA AGATCTGATC
TTGTCCAGCC CTGACTATGC TAATAATATC ACCACATACA TCAATGCCGA CAAGGCTCGT
ATGAGCGGTA TCGAGGCCCT TTTGTCTTAT GACTTCGGCA GCCTCTTTGC CAACAAGTTC
                                                                                               1560
                                                                                               1620
                                                                                               1680
        TCTCTCCGCG CATTTGCGAA TGCCACGATC ATGCTCAATT CCGAGATGAA GAAAAGCCAG
                                                                                               1740
       ACCGATGCCC CTTGGAGCGA AATGTACTAC GTTCGCAAGC AGAACATCAC CTTCGGTATC
                                                                                               1800
60
       GAATATCGTG GCAAAGAAGG ACTTGAAGTG ATGCTCAACG GTCGCTTCAT GGGACGCAGG
ATCGAGCAAA ACTGGTATGC TTACTACCCC GAAGTTCGCC CCGAACTCCA GCAACTGCTT
                                                                                               1860
                                                                                               1920
       GCAGCAGAAG AGCCTGAATT GGCTGCTCAG GGACTGCTCC GTCATCCGCA AGCAATGGTG
                                                                                               1980
       TTCAATGCCT CTGCTTACTA CCACATGAAC AAGTATCTCA CCTTCGGTGT GAACTTGAAC
                                                                                               2040
       AACATCTTGG ATGAGCTTTA TACGGAGAAA GACGGCTACC ACATGCCCGG ACGTAACATC
                                                                                               2100
65
       ATGGGTAAGG TTATGGTCAA CTTC
                                                                                               2124
       (2) INFORMATION FOR SEQ ID NO:185
70
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 1386 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
                    (D) TOPOLOGY: circular
75
```

WO 99/29870 PCT/AU98/01023

```
(ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
 5
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
10
          (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1386
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:185
15
      ATGAACAGGT TTTCAAATCA TTGGCCCTGC ATCCTCGTGG GGTTTGTACT CTGGTTTGTA
                                                                                60
      TCGGCGAGTC GGACTGTGGC ACAAAACGCC TCCGAAACGA CGGTATCGTA CGATACGGAT
                                                                               120
      ACCGCCGTAC TCTCCGAAGC CGATGTGCTT CGGATCGCTC TTAGTGAGAA TGCCACAGTG AAAGTGGCCG ATATGGATGT GCGCAAACAG GAATATGCAC GTAGGGCAGC ACGTGCCGAT
                                                                               180
                                                                               240
20
      CTCTTCCCGA AAGTAGACCT CAATGGCGTT TACAGCCATA CGCTAAAGAA GCAGGTCTTA
                                                                               300
      TATATAGATA TGCCCGGTTT CAGCAGTAGC GAAGGTATCG AAATGGGGCG TACACACAAT
                                                                               360
                                                                               420
      ACGCAAGGAG GGGTGAACGT CTCCATGCCA TTGGTGTCGG CACAGCTTTG GAAAAGCATT
      SCCATGACCG GAGAACAGCT CGATCTGGCT CTGGAGAAAG CTCGCAGCTC CCGAATCGAT
      TTGGTGGCAG AGGTGAAGAA GGCTTACCTC AGTGTATTGT TGGCCGAGGA CTCTTATGGC
                                                                               540
25
      GTATTCAAGC GCAGCTATGA CAATGCTCTG GCCAATTATA AGAACATATC CGACAAGTTC
                                                                               600
      GATCGTGGAC TTGTGGCCGA GTATGATAAG ATTCGAGCCA ATGTACAGGT ACGCAACATC
                                                                               660
      GAGCCTAACC TCTTGCAAGC GCAGAACTCC GTAGCCCTTG CTCTCTGGCA GCTCAAGGTC
                                                                               720
      CTGATGAGCA TGGAAGTGGA AACTCCGATC AGACTCTCCG GTTCATTGTC CGACTATAAA
                                                                               780
      GAACAAGTOT ATACCGGCTA TITTGCCGCC GATACGCTTA TITCCAACAA CTCCTCCCTG
                                                                               840
30
      CGTCAGCTCG ATATACAGCG TCGTCTGGCT GTCAGTGCAG ACAAGCTGAA CAAGTACAGC
                                                                               900
      TTCCTGCCTA CACTCAATCT GGGAGGGCAG TACACCTATT CGCTCAACAG CAACGACATC
                                                                               960
      AAATTCTGGG GCGAGGGACA ACGCTGGACG CCTTTCTCCA CCATATCGCT CAGCCTGTAC
                                                                              1020
      ATTCCTATAT TCAATGGAGG CAAACGTCTG TACAACGTGA AGCAAAGTGC TTTATCGATC
                                                                              1080
      CGTCAGATCG ATCTGCAACG ACGCCACATA GAGCAATCCA TCCGAATGGG AATCAAGAAC
                                                                              1140
35
      CAAAATGACC GTCTGCGTAC CTGTATGCAG AGATTTGTGG CCTCGGAAGA GGCTGTCCGA
                                                                              1200
      AGTGCAGAAA AGGGCTATCA GATAGCAGAG AAACGCTATC AGACAGGCGA AGGCACTCTC
                                                                              1260
      GTCGAGCTCA ACGATGCCGA TGTGGCTCTT TTGCAGGCTC GACTCAATTA TAATCAGGCC
                                                                              1320
      ATATTCGACT TTATGACCGC AAAGGCCGAA TTGGACAAGA TGAACGGCAT GGGGATTCCC
                                                                              1380
                                                                              1386
      GAACAA
40
      (2) INFORMATION FOR SEQ ID NO:186
           (i) SEQUENCE CHARACTERISTICS:
45
                 (A) LENGTH: 1476 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
50
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) AUTI-SENSE: NO
55
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
60
                (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1476
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:186
65
      ATGTGGGGG ACAGCCATGG AGTGGCGCCG AACCAAGTGC GCCGAACGCT GGTGAAGGTA
      GCCTTAAGTG AATCCCTTCC TCCGGGTGCA AAACAGATTC GTATCGGATT CTCTCTTCCG
                                                                               120
      AAAGAAACGG AGGAAAAAGT CACCGCCCTA TATCTCCTTG TGAGTGATTC TTTAGCGGTG
                                                                               180
      CGCGACTTGC CGGACTACAA AGGGCGAGTC TCTTACGATA GCTTCCCGAT CTCAAAGGAA
                                                                               240
      GATCGTACCA CAGCCCTTTC TGCGGATTCG GTAGCCGGAC GCCGCTTCTT TTATTTGGCT
                                                                               300
70
      GCGGATATAG GGCCTGTTGC TTCTTTTTCC CGATCCGATA CGCTGACTGC CCGTGTGGAA
                                                                               360
      GAGGTGGCTG TCGATGGCCG CCCTTTGCCG TTGAAAGAGC TGTCGCCTGC CTCCCGTCGT
                                                                               420
      CTGTATAGGG GGTATGAGGC CCTCTTTGTA CCCGGTGATG GCGGATCGCG GAACTATCGT
                                                                               480
      ATCCCGGCCA TTTTGAAAAC GGCTAATGGA ACACTCATAG CGATGGCCGA CAGACGAAAA
                                                                               540
      TATAATCAGA CGGATCTGCC GGAGGATATA GATATAGTCA TGCGGCGCAG TACGGACGGA
                                                                               600
75
      GGGAAATCGT GGAGCGATCC CAGGATTATC GTACAGGGAG AGGGGCGCAA TCATGGCTTT
                                                                               660
```

```
GGCGATGTAG CCCTGGTGCA AACCCAAGCA GGAAAGCTCC TGATGATCTT TGTCGGTGGA
       GTAGGCCTGT GGCAGTCTAC CCCCGATCGT CCTCAGCGCA CTTATATATC GGAAAGTCGG
                                                                               780
       GACGAAGGAC TGACTTGGTC GCCTCCTCGG GATATAACCC ATTTCATCTT CGGCAAGGAT
                                                                               840
       TGTGCCGATC CGGGACGCAG TCGCTGGTTG GCCTCCTTTT GTGCTTCGGG ACAAGGGCTT
                                                                               900
 5
       GTGCTGCCAT CCGGTCGTAT CACGTTTGTG GCTGCCATCC GCGAATCAGG GCAGGAGTAC
                                                                               960
       GTCCTGAACA ACTATGTCCT CTATAGCGAC GATGAGGGCG ATACATGGCA GCTTTCCGAC
                                                                              1020
       TGTGCATACC GCCGTGGCGA TGAGGCAAAG CTTTCATTGA TGCCCGATGG CAGGGTACTG
                                                                              1080
       ATGAGCATAC GCAATCAGGG ACGGCAGGAG AGCCGACAGC GTTTCTTCGC TCTCTCCTCC
                                                                              1140
       GACGATGGCC TTACTTGGGA GAGAGCCAAG CAGTTCGAGG GCATCCATGA CCCCGGCTGT
                                                                              1200
10
      AATGGAGCTA TGCTTCAAGT GAAAAGGAAC GGAAGGGATC AAGTGCTGCA CTCCCTGCCT
                                                                              1260
       CTCGGCCCGG ATGGGCGTCG CGATGGAGCT GTCTATCTCT TCGATCATGT CTCCGGCCGC
                                                                              1320
       TGGTCCGCTC CCGTTGTTGT CAATTCAGGA TCGAGTGCCT ACTCGGATAT GACTCTGCTG
                                                                              1380
      GCGGATGGAA CGATCGGTTA TTTCGTCGAA GAGGGCGATG AGATCTCATT GGTTTTCATT CGGTTCGTCC TTGACGATCT CTTCGATGTC CGGCAA
                                                                              1440
                                                                              1476
15
       (2) INFORMATION FOR SEQ ID NO:187
            (i) SEQUENCE CHARACTERISTICS:
20
                 (A) LENGTH: 735 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
25
           (ii) NOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
30
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
35
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...735
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:187
40
      ATGAAAAAG AAAAACTTTG GATTGCGATC GTCGCCGGTT TGGCTTTCGT ATTGGGCCTT
      TATGCTCTTG GCCGCAGTGT CGCTCAGCTA CGCCGCTCTC AGCCTTCGGT GACTGTGACC
                                                                               120
      GGTATGCCG AGCGTAATTT CAAATCCGAT CTGATCGTTT GGACTGCTTC GTACCAGCTC
       CAGATGATGG ATCTCGAATC GGCCTACAAG GCTTTGAAGG AAAAACAGAT ATTGGTAGCA
                                                                               240
       GACTATTTGA AAAACAAGCA GCTGCCCGAT TCGTCTTATA TCTTCTCAAG CGTAGCCATC
                                                                               300
45
      TCTAAAGAAT ACAACTACTA TTACGATCCT CGGCAGGAAC AAAACGTCAG GACCTTTGCC
                                                                               360
       GGGTATCTGC TCAGCCAGAC AGTTACGGTG ACCTCACAGG ACATCGAACA TGTGGAGAAA
                                                                               420
       ATATCTCGCG ATATAACGGA GCTGATCAAT CAGGGGGGTAG AGATTACCTC CGACCGTCCG
                                                                               480
       GCCTATTACT ACACCAAGCT CAATGATCTG AAGGTGGAGA TGCTGCGCAA TGCCTCCGAA
                                                                               540
      GACGCTTTCA ATCGTGCTTC GGTCATTGCG GAGGGGAGCG GTTCCTCCGT GGGTAAGATG
                                                                               600
50
      CTATCTTCTT CGATGGGCGT GTTCCAGATA GTGGGGCTCA ACTCGAACGA AGATTATAGC
                                                                               660
       TGGGGAGGTT CGTTCAATAC GTCTTCCAAG ATGAAGACGG CAAGCATAAC GGTTAAGGCT
                                                                               720
      TCTTTCGCTT TGAAG
                                                                               735
55
       (2) INFORMATION FOR SEQ ID NO:188
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 828 base pairs
                 (B) TYPE: nucleic acid
60
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) NOLECULE TYPE: DNA (genomic)
65
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
70
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...828
75
```

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188
       ATGAAAAAA CAATTGCAAT TATCGCCTCA GCCCTCTTGG CTTTAGGAGC CGTCGGCTGT
       AAGAAAAATG CTGACACTAC CGCTGTCAGT GAAAAGGATA GCATAGCCTT GTCCATGGGT
                                                                                  120
 5
       ATTTTGTACG GACAGGATTT TGCCAATCAG TTCGAAATGT CCCGCTTGCA AGGCCAGCCG
                                                                                  180
       ATTGATTCGG TAGCTTTCTT GGACGGTTTC AAATATGGTA TCGATACGAC GCGCTTCTCG
                                                                                  240
       TACAATCTGG GAGCCATCTA TGCTTCCAAT ATAGCTCGTC AGCTGGCTCA TGATTCCATC
                                                                                  300
       GATATCGACA AGTTCTATGC AGCCATGCGT GCGGCTCTTC TTAAAGACAC CGTATCTATC
                                                                                  360
       GCCATGAAGC CTGCAGATGC ACAGGCTTTC ATGCAACGAA TCCAAGCCAA AAAGCAGCGA
                                                                                  420
10
       GAAAACAATA TGAAGCAGTT TGGCCAGAAC ATCGAAAAGG GTAATGAATA CATCGATACC
                                                                                  480
       TTTAAAAAAG AAGATGGTGT AACTGTTACG ACAACTGGTC TGGCATACAA GACTCTTCAG
                                                                                  540
       GAAGGTACGG GAGCTACTCC CTCTTTGGCC GATACTGTAC GTGTCAAGTA TGTGGGTACT
       CTGGTCGATG GTAAAGAGTT CGACAAAAAC GAAGAAGGAA TCGAATTTGC CGTTACCGGT
                                                                                  660
       GTGATTAAAG GCTGGACGGA GATGCTCCAA CTCATGAAGG TCGGTCAGAA AGTTCGCGTG
                                                                                  720
15
       GTAATCCCAC AGGAGCTGGC TTATGGGGAG ACCGGCAACT ATACCATCGA ACCGTTCTCT
                                                                                  780
       ACCCTGACGT TCGAGATGGA ACTTATCGGG ATCAAGCCCG GGAAAAAG
                                                                                  828
       (2) INFORMATION FOR SEO ID NO:189
20
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2325 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
25
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
30
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
35
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2325
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:189
       ATGAAAGTAT TACGGCAAGT ATTCCTCCCC ATCCTTTTG TCCTACTGAC AGGTGCCTGC
                                                                                  60
       TCCACCACAA AGAATCTGCC GGAAGGCGAA CAGCTGTATA TCGGAATGGG CAAGACACAG
                                                                                 120
       ATACTCCGGC AGGACAAGAG CCACGCCGGC CAACAGGCTC TGACCGAAGT GGAGAGTACA
                                                                                 180
45
       CTGAAAGTTA CACCCAATGG AGCTATTTTC GGCAGTGCAA GTGCCTCCTT ACCCAAGATA
                                                                                 240
      CCATTCGGGC TATGGCTATA CAACAGCTTC GTGGGGGATT CCACTGTCAT TTCGAAATGG
ATATTCGACA AGTTTGCAGC CAAGCCGGTT TTCATCAGTC AGGTCAAATC CGATAGCCGG
                                                                                 300
                                                                                 360
       GCTAAGGTGG CGACGAACAT CCTCCGCGAA CACGGGTACT TCGATGCTAA AGTAAAAAGC
                                                                                  420
       AGTGTGACCA CTCTGAAAAA GGACTCGCTC AAAGCCAAAA TCTCCTATAC GGTGGATATG
                                                                                 480
50
       GCCTCTCCTT ATCATTACGA CAGCATCATT CCCTTACCGA TCAGCACTTT CCCCGACAGC
                                                                                 540
       ATTCTGGCTT ACAGGCAGAC TCCGTCTTTG ATCAGGAAAG GAGACCAGTT CAATTTGGCA
                                                                                 600
       AAGCTGCACG AAGAGCGTCA GACCATCAGT GCCCTGCTGA GAGACAATGG TTACTACTAC
                                                                                 660
       TTCCGCCCAC AGGATATTAT CTACGAAGCC GATACCCTCC TCGTAAGAGG TGCCGTATGC
                                                                                 720
       CTGCGAGCCA AGCTCTCGGA AGATACTCCA CCCCAAGCCA TGCGCCCGTG GAGGATAGGG
                                                                                 780
55
      AAACGGACAG CAGTCCTGCT CGGAATGAAC GGAGAAAGCC CGACAGACTC GCTCGAAGTG
                                                                                 840
       GAGGATATGA AAGTCCTTTA CTATCGTAAA ATGCCGGTTC GCCCCAAGAT TTTGGCCAAA
                                                                                 900
      CGCTTTCGTT TCTTCTCCGG CAATCTGTAT CGGCAGAAAG ACGATGAGAC GACACGCAAA
TCCTTGGCTC GTTTGGGAGC CTTCTCCGTT ATCGATCTCA ATTTTTTGCA ACGCGATTCC
                                                                                 960
                                                                                1020
      ATTTCCGGCC TTTTGGATGT GCGACTGCTA ACCACCTCG ACAAACCTTG GGATGCATCA TTAGAGACCT TGTTCACGAG CAAAAGCAAT GACTTCATCG GTCCCGGACT GAATTTGCT
                                                                                1080
60
                                                                                1140
       CTTGCTCGGC GCAATGTATT CGGCGGAGGA GAAAATCTTT CTTGGAATAT CGGTGGATCG
                                                                                1200
      TATGAGTGSG AGACCGGCAA TCGTCCCGAA AATAGCAGCA ATCGGCTGAT CGATATAAAT
                                                                                1260
      TOGTACAACA TGAATACGGO CGTGAACCTC TOGTTTCCCT CGATTGTATT TCCCGGTCTG
                                                                                1320
      CTGGATAAAT ACTATTACTA CCCCACGACT ACGACTTTTC AGGCTTCTGC CACCGCGCTG
                                                                                1380
65
      AACAGGGCAC ACTACTTTAG CATGTACTCT TTCGGCTTTT CGACCACCTA CGAATTTCAG
                                                                                1440
      CCCTCCAAGG AACACCGGCA TGCTATTTC CCGCTCAAGC TCAACTACAA CCTCCTGGGG
                                                                                1500
      CATCAGACAG AAACTTTCCA GGCCATTACG GCGAACAATC CGCCCCTGCT GCTCAGCCTT
                                                                                1560
      CAGAGTCAGT TCCTTGCTCA AATGGGGTAT ATCTATACGT TCAACAAATC CGTTTCAGAG
                                                                                1620
      AAAAGTCCTC ATCATCTTTG GATGCAATTC GGACTATCCG AGGCAGGCAA TCTCCTGAAT
                                                                                1680
70
      CTGATCTATC TGGCAGCCGG CAAGAAGTAC AGCGACACCA AGAATTTCGT CGGCGTCCCC
                                                                                1740
      TTCTCTCAGT TCATCAAAGC CACGGGAGAA CTGCGCTATT CCTATACCAT AGACCGCAAT
                                                                                1800
      CAGTCACTGG CAACCCGTTT CGGGACAGGC GTGATATATA GCTATGGCAA TATGCGAGTG
                                                                                1860
      GCACCCTATA GCGAGCAGTT CTATGTAGGC GGTGCCAATA GTATCAGAGC TTTCACCGTC
                                                                                1920
      CGTAGCATCG GCCCCGGACG GTTCAATCCG GATTCCGACA ATCAGTATTC CTATTTGGAT
                                                                                1980
75
      CAGGTGGGCG AATTCAAACT CGAAGCCAAC GTGGAATATA GAGGCAAGCT TTTCGGGGAT
                                                                                204û
```

```
CTCCACGCAG CCGTTTTCCT CGATGCGGGC AACGTTTGGC TCTTGAGGGA GGATTCTTCC
                                                                             2100
      CGTCCGGGCG GTGCTCTGTC CGAAGTGGGA TCGGTGAGCA ATTTCCTGAA TAGCATCGCT
                                                                             2160
      CTCGGCACCG GTGTCGGCCT TCGCTACGAT CTGGCATTTC TCGTGGTTCG TGTCGATGTC
                                                                             2220
      GGCTTCGGTC TCCACCTTCC TTACAATACG GGTAAGAAAG GTTACTACAA TATCCCACGC
                                                                             2280
 5
      TTTAAGGATG CCATCGGTTT CCATTTGGCT GTCGGCTATC CCTTC
                                                                             2325
       (2) INFORMATION FOR SEQ ID NO:190
10
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2322 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
15
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
20
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
25
          (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...2\overline{3}22
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:190
30
      ATGTCCTCGC ATTCCGTTCG GTATCTAATC GGCATTGCCG GCTGCTTGCT CCTCATGCTT
      GCTTCCTCCT GCTCGGTCAC CCGTTATGTG CCGGACGGTA GCAGACTATT AGACAGGGTA
                                                                              120
      ACGATCGCAA GCGAAACGGG CAGTATCGCT CTGCCGGAAG ATATTCGGGA CTATACCCTC
                                                                              180
      CAGCAACCCA ATTACAGACT GTTCGGGATG ACTCGCTGGC TACTGCGCGT CTATAGCAGC
                                                                              240
35
      TCGAATCCGA ACAGCAACAG CTGGTGGAAC CGTTCGCTCC GGAAAATGGG CGAACCGCCT
                                                                              300
      GTCCTCATCG ATTCTGTCCT CACCGATCGT ACTGCCAACC GTCTGGCAAA GGCGATGGCC
                                                                              360
      GGCGATGGCT TTCTCGATGC TACTGCTCGT GCCGTGGTAG ACACCGGCTT GTACAAGAAA
                                                                              420
      GCTCGCATTA CTTATCTGAT TCAGCCCGGA AGCCGTTATT ATATACGCAA TATGGCTTTG
                                                                              480
      GATGTGAAGA ATCCACTCCT TCCTCCGTT GCGCTTGGCA ATTCGCTTCC TTCGGCATAC
                                                                              540
40
      AAGGTCGGGA TCAGCGAGGG TTCTCCCTTG TCGCCCATCG TACTCGATGA AGAGAGAAAG
                                                                              600
      GCGATAGCTC GTCATATGCG CAACAACGGC TTCTGGAAGT TCTCCGCCGA GGATGTTTAT
                                                                              660
      TATGAAGCAG ATACTACCGT TTCAGGAGGA TCGGGTACGA AATCTGCCGA TCTGAAATTA
                                                                              720
      GTGGTCAATG GCATCGGGCG TTATCCATAT CGGATCGGCA GGGTATTCTT TCATGCCGAT
                                                                              780
      TATGATCCTC TCGAATCGGA CTTCAGAGTT CAGGAGCTGC CACGTATCGA TTCGATTTCG
                                                                              840
45
      CGTGGCGATT ACACTGTTTA CTATGGGAGT AGGGGACGTT ATATCCGGGC ATCGGCTCTC
                                                                              900
      ACGCGGTCGG TGTCCGTTAC ACCGGGAGCT TTTTTCTGCG AGGATGATGT GGAACGCTCT
      TATATCAAGC TGAATGCGCT CCCTATCGTT CGGAACGTGA ATATCCGATT TGTGGAGCAC
                                                                             1020
      AATGGTAAGG ATGAGATTGC TCTGGCGGAT AGCTCTCGCC TTGTGGACTG CTATATTCTT
                                                                             1080
      ACCGTTCCGG CCAAGAGCAA ATCGTTCGAA GCCGAAGTCC TCGGCACCAA TTCCGCTGGA
                                                                             1140
50
      GACTTCGGGG CGGCTTTGTC TCTCGGTTTC ACCGATCGCA ATTTGTTTCG TGGGGCGGAG
                                                                             1200
      ATGTTCAATA TCAAACTCAA GGGTGCTTAC GAAGCCATTC GCAAGGGTTC GCACAGCTTC
                                                                             1260
      ATGGAATATG GGGTGGAAAG CTCGCTCCGT TTCCCTCGTC TCCTCTTCCC ATTCATTTCT
                                                                             1320
      GACGAAACGC GCCGGCGGCT ACGGGCATCC ACGGAATGGA AGATCGGGTA TAATTACCAG
                                                                             1380
      ACACGTCCGG AGTTTGATCG GGTGATTCTC TCCGCTCAAC TCAATTATTC ATGGCAGACC
                                                                             1440
55
      TACCTGCACA ATCGTCTGCG TCATACGATC CGCCTGCTGG ATGTCGATTA TCTCCATCTC
                                                                             1500
      CCGTACATCG ATCCCGACTT CGCCCAATCC CTTCCGCCTA CGACTGCACT GTATAACTAC
                                                                             1560
      ACGGAGCAGT TTATCCTCGG CTCGGCATAT ATACTGAACT ATACCACGGC TTCGTCCATG
                                                                             1620
      GAGCGTACCG TATCCAATCC TTTTACGGCA CGGTTCAGTA TCCAGACAGC CGGCAACCTG
                                                                             1680
      CTGCAAGCCA TTTCTTATCT GACCGATTCT CCGAAAGACG AACACGGGTT GTATAAAATG
                                                                             1740
60
      TTCGGTCTGC ACTATGCTCA GTTCGTCAAG CTCGATCTCG ATCTGGCTAA AACCGTTCTT
                                                                             1800
      CTCGAAAAGG ACAATACTTT GGCACTGCAT CTGGGTTTCG GACTGGCTTT CCCTTATGGC
                                                                             1860
      AATGCTCGCC ATATACCCTT TGAGTTACGT TACTTTGCCG GAGGATCGAA CAGCGTTCGC
                                                                             1920
      GGCTGGAGTG TCCGTACCCT CGGCCCGGGG AGTATGAAGA TGACTCCGGA CAAGACCTTC
                                                                             1980
      TTCGATCAGA TGGGTGATAT TCGTCTGGAT CTGAATGTCG AATACAGGAC AAAGCTGTTC
                                                                             2040
65
      TGGAAGTTTC GCGCAGCAGC TTTTGTCGAT GCCGGCAATG TCTGGACGAT AAAGGAGTAT
                                                                             2100
      GAGAATCAGG AGGACGGTCT CTTTCGTTTC GATCGCTTCT ACAAGGAAAT AGCTTTGGCC
                                                                             2160
      TACGGTCTGG GGCTTCGTCT CGACTTCGAT TATTTCCTTG TGCGGCTGGA TGCCGGACTG
                                                                             2220
      AAAGCCTACG ATCCTCAGCA GACAGGGCGT TACAAATGGG CTATCACACG CCCAAACCTT
                                                                             2280
      TCTTCCAATT TCGCTTGGCA CATTGCAGTA GGCTATCCGT TC
                                                                             2322
70
      (2) INFORMATION FOR SEQ ID NO:191
```

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2601 base pairs

75

```
(B) TYPE: nucleic acid
                   (C) STRAHDEDNESS: double
                   (D) TOPOLOGY: circular
 5
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
10
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
15
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2601
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:191
20
       ATGAGAAAAA GAATTCTACA ACTTTTCCTG ACCGCATTGC TGCTGGCATT AGGCTCCTCT
       CTCGCCATAG CGCAAACAGT GGTGACCGGT AAGGTGATCG ATTCAGAAAC GTCCGAACCG
                                                                                     120
       CTCATCGGTG TATCCGTAAG CACCGGTCAG GGAGCATCCC TCCGCGGTGT AACCACCGAT
                                                                                     180
       ATGGATGGTG GCTTCCGATT CGAAGTACCG GCCAAATCTG TCTTGACTTT CCGTTGCGTA
                                                                                     240
       GGTTATGCTA CCGTAACTCG CTCTATAGGC AGAGGTTCTC AAGAAGACCT CGGTACGATT
                                                                                     300
25
       CTCCTCGATC CCCAGGCCAT CGGCTTGGAT GAGATTCAGG TAATAGCCTC TGTGGTGCCC AAAGACCGTA TGACGCCGGT ACCCGTTTCC AATATCCGTG TGGCTGATAT TCAGGCAGCA
                                                                                     360
                                                                                     420
       TCGTTGAATG TCGAATTTCC CGAACTGGTT AAATCCACTC CCTCTACCTA TACGACAAAA
GGAAGCGGAG GTTTCGGTGA TGGTCGTACC AATGTGCGTG GATTCGACAC TTACAACTTC
                                                                                     480
                                                                                     540
       GGTGTACTCA TCAACGGAGT TCCTGTCAAT GGTATGGAAG ACGGGAAAGT ATATTGGAGC
                                                                                     600
30
       AATTGGAGTG GTCTGATGAA TCAAGCCAGT ACCATTCAGA TTCAGCGCGG ACTCGGAGCC
                                                                                     660
       TCCAAGCTCG GTATCAGCTC GGTAGGTGGT ACGATGAACA TTATCACGAA GACTACGGAC GCCAACACCG GAGGTTCGGC TTATGTCGGT ATGGGTAATG ATGGATTGCA CAAAGAATCG
                                                                                     720
                                                                                     780
       TTCTCCATTT CTACGGGTAT GAACGACGGT TGGGCTATCA CCATTGCAGG CTCCCATATG
                                                                                     840
       ACGGGTCTGG GTTATGTGAA GGGGCTGAAG GGACGTGCAT TCTCTTACTT CTTCAACGTT
                                                                                     900
35
       TCGAAGAAGT TCAATGAACG TCATACCCTC TCTCTTACCG GATTCGGTGC ACCACAATGG
                                                                                     960
       CACAACCAAC GTTCTTCCAA ATATTCTGTA GCCGACTATG ACAAATACGG CATCCGTCAC
                                                                                    1020
       AATCAATCCT TCGGCTATCT GCGAGGCGAA CTGACTCCTA CGGCTTATGC TTACAATACG
                                                                                    1080
       TACCACAAGC CCCAGTTCTC GCTGAACCAC TTCTGGAAGA TGGATGAAAA TACCTCTCTT
                                                                                    1140
       TATACOGCAN CCTACGCATC TTTGGCTACC GGTGGAGGTC GTCGCGCTTA TGGAAAGAAC
                                                                                    1200
40
       AGTAAGTGGG TATTGATCAA CTACAACACC GGACAACCCT ATGAACAAAC AAAGGTGACT
                                                                                    1260
       CCCGATGGAC TTATCGACTA CGATGCCGTA CTGGCTGCCA ATGCTGCGGC GAGCAATGGC
                                                                                    1320
       TCGGAAGCAA TTTTTGCCCT TGGCTCCAAC TCTCACAAGT GGTTCGGTCT ACTCTCTTCA
TTCAAGAAGA AACTTAATAG TTCGCTGACT TTGACAGCCG GATACGATGG GCGTTACTAC
                                                                                    1380
                                                                                    1440
       CGTGGCGACC ACTATGACAA GATCACCGAT CTGCTCGGCG GTAGCTACTA CATAGAGGAT
                                                                                    1500
45
       CCCAAGACAA AGCTCGCATA CCATGCGGAA GGTCAGCAAC TGAAAGTGGG TGACATTGTA
                                                                                    1560
       AATCGGGACT ACACAGGCGA AATCATGTGG CACGGCCTCT TCGCACAGAT GGAGCATTCG
       TCCGAATGGA TCGATGCATT CGTATCAGGA TCTATCAACT ACGAACTATA CCGCAATCAC
                                                                                    1680
       AACTATGGCG GTAGCAAGTC CACCGGCTAC CTGCCCGGCG TATCGCCGTG GAAAAGCTTC
       CTTCCGTGGA GTGGCAAGGC AGGTCTGAGC TACAAGTTCG CACAGGGACA CAATGTATTC
                                                                                    1800
50
       GCCAATGGCG GTTTCTTCAC ACGTGCACCA CTCTTTGGCA ATATCTATGC TGCGGGGGCT
                                                                                    1860
       ATCATTCCCA ATGACAAAGC CAATATGGAA AAGGTGCTTA CAGGAGAGGT CGGCTATGGA
                                                                                    1920
       TTCACGAATC ACAAAAACTT CGAGTTCAAT ATCAACGGAT ACTATACGAA GTGGATGGAT
                                                                                    1980
       CGCGTGACCT CGAAGAGAAT CGGAAACGAG TATGTTTATC TCAATGGCGT TGATGCTGTT
                                                                                    2040
       CACTGTGGGG TAGAGGCTGA GGTCAGCTAT CGTCCTATTC GTCAGATCGA CCTTCGCGGT
                                                                                    2100
55
       ATGTTCTCTC TCGGTGACTG GACTTGGCAA AACAATGTAA GTTACACTTC TTACGACGAA
                                                                                    2160
       GCCGGCAATG AGACAGGGCA GGATATAACC TATATCAAGG GTCTTCACGT CGGAGATGCA
                                                                                    2220
       GCACAGATGA CGGCTGCTGT ATCGGCAGAC ATAGAGCTGT TCAAGGGTTT CCATGTCATA
                                                                                    2280
       GGTAAGTACA ACTTCCTTGG CAAGAACTAT GCAGGATTCA ACCCCGCAAC GCGTAATGCA
                                                                                    2340
       CAGCAGTACG AAGCGGATGG CAAAGAAATC GTGGAATCAT GGAAGTTGCC CGATGTAGGT
                                                                                    2400
60
       CTGTTCGATC TGTCTGCATC CTACAATTTC AAGCTTGGTT CACTCAGCAC CACATTCTAT
                                                                                    2460
      TTCAACATGG ACAACGTAGC CGACAAGCGA TATGTGAGCG ATGCCGACGA CAATATCATC GGTAAGAAAC ACGATGAGGC TTCGGCTCTC GTATGGTACG GTTTCGGCCG CACTTGGTCT
                                                                                    2520
                                                                                    2580
       ACCGGTATTC GTGTAAACTT C
                                                                                    2601
65
       (2) INFORMATION FOR SEQ ID NO:192
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1293 base pairs
70
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
75
```

```
(iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
 5
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
10
                 (B) LOCATION 1...1293
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:192
      ATGAAGTTTT CAATCCGCCT TTTCCTCTGC ATCATCTTTC TCCTCTGC ATTTATCCTG
15
      CCTGCTCTCG GACAAAAATC CAAGCAGGTA CAGCGACTTG AGAAGCAACG TAAGGAGGCC
                                                                                 120
      CTCAAAGCCA TCGAAAAAAC CGATCGCGAA CTACGAAATA CCAAGAAAGA CAAGCAAGAC
                                                                                 180
      AAACAAAAGC ATCTCAACCT CCTGAACAAG CAGGTTGCTC AACGCAAGCA GATGGTACAA
                                                                                 240
      CTCTTGGACA ATGAGGTCAA AGAGTTGCAA TCCGACATTG ATTCCATGAC GGGTGTATGT
                                                                                 300
      CATCAGCTCT CTGTAGAAGA GAAAGCCCGA TCCGATGAAT ATGCCCAAGC TCTACAGTCT
                                                                                 360
20
      ATGCAAAAGC GGAAACGCTC GTTGGATCGC ATCCTTTTCA TTTCATCGGC CAAGAGCTTT
                                                                                 420
      GACGAAGGCA TGCGACGGAT GCGTTTCTTG GAACAATACG CTTCTGCATA CAAGCTGGCA
                                                                                 480
      TCTGTCCGGC TGCGCGATAC ACGTAGCAAG TTGGAGACTG AACGTGCGAC TGTAGAAGAC
                                                                                 540
      GCCAAAAAGG AGAAAGGACA TCTCTTAGTC ATCAGAGAAG AGGAAAAAAA GAAACTCGAA
      GGACAGCAAG CCGAGCAACG TCGGCAGGTG CAGGCTTTGG GAGCCAAACA AAAAGACTTG
                                                                                 660
25
      GAASCGCAGC TGCGAAAGCA GAAAAAGCAA GCCGAAGCTC TGAACAGAAA GATCGAGAAA
                                                                                 720
      CAGATTGCCA AGGAAATAGA AGCTGCCGAA CGTCGTGCTC GAGAAGAACG TGAACGGTTG
                                                                                 780
      GCACGCGAAG CCAAAGCCAA GGGTAAGCCG GTTCCTGCCG AACCGGAACG GAAGGCGGAG
                                                                                 840
      ACCAAAGGCG GCTATGCTAT GGATGCCTCT GAGCGTGCTC TCTCGGGCAG CTTTGCACAG
                                                                                 900
      AACAAAGGTC GCCTGCCCGG CCCCGTTCGC GGCAGATACC GAATCGTAAG CGACTTTGGC
                                                                                 960
30
      GTGCATCAGC ACAGTGAGCT GAAAAAAGTA CAAGTTAATA ATGGAGGTAT CGACATCGCT
                                                                                1020
      GTAGCAACAG GATCCGATGC TACCAGCGTA TTCGATGGTG TAGTGTCCAG TGTATTCGTG
                                                                                1080
      ATACCCGGTT ATAATTCGGC CGTAATGGTT CGTCACGGTA ACTATATCAC GGTTTATGCG
AATCTGAGCA AAGTGTATGT AAATTCCGGC ACTCGTGTTA AAACGGGTCA GGCTCTTGGT
                                                                                1140
                                                                                1200
      CGTGCCTATA CGGATCCTTC CAACAACCAG ACCATTATTC ACTTCGAAAT CTGGAAAGAA
                                                                                1260
35
      CGCAGCAAAC AAAACCCAAG ACTATGGTTA CGA
                                                                                1293
      (2) INFORMATION FOR SEQ ID NO:193
40
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 999 base pairs
                 (B) TYPE: nucleic acid
                    STRANDEDNESS: double
                 (D) TOPOLOGY: circular
45
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
50
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
55
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...999
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:193
60
      ATGAAAAAGT ATTTGTTATA TGCCTCGTTG CTAACGAGTG TTTTGCTCTT TTCCTGTTCA
      AAGAACAATC CTAACGAGCC GGTGGAAGAC AGATCCATCG AAATTTCTAT AAGGGTAGAT
                                                                                 120
      GATTTCACCA AAACGGGTGA GGCAGTACGC TATGAAAGGA ATCAAGGAAG TGCTGCCGAA
                                                                                 180
      AGGCTCATTA CCAATCTTTA CCTCTTGTTG TTCGATCAGT CAGGGGCGAA TCCGGCGAAA
                                                                                 240
65
      TACTATATTA CCGGTAACAC TTTCACCGGA GGGACCTGGC TTCCTGACGA TATGAAGGTG
                                                                                 300
      AAGTTGGATA TGACACAATC CGAGGCCGGA GAGCGCAAAG TATATGTCGT AGCCAATGTT
                                                                                 360
      GATAATGCGG TTAAAACGGC TCTTGATGCT GTCGCTAACG AAAGCGATTT GCAGACTGTA
                                                                                 420
      AAGAGGACGA CTGCAATGCC GTGGTCGACC GATATAGCCT CTCCTTTCCT GATGTCCGGA
                                                                                 480
      AACAAGACAC ACGACTTCTT GGCCAATCGT CTTTTGGACA ATGTGCCCCT TGTGCGTGCC
                                                                                 540
70
      ATTGCCAAGG TGGAGCTGAA TATCTCGCTG AGTGAGAAAT TTCAGATTGT GCCGATAATT
                                                                                 600
      GTCAATGGTA GTTTGAGTGA GTTCAAGTTC AGATACGTAA ACTTCGACAA GGAGACCTAC
                                                                                 660
      GTAGTGAAGC CAACGACCAA GCCGGACAAT CTCATTAGTT CTGCTAATGG TGTTTGGCCT CAGATTACAG ATTGGACTGT ATGGGGTGCT TCCTTAAATA CTTCTCCTGC TCCGGATGCG
                                                                                 720
                                                                                 780
      GGCACAGGTT ATACATTGGA TGCAAATGGC AAGGTAACGG CACTACGGAT TGTTACCTAT
                                                                                 840
75
      CTGAATGAGC GCGATAGCAA AGGGGCTACG GTAGAGGTCG CATTGCCTCG TGTGGATGAT
                                                                                 900
```

| | GGCACCCTTC CTCCTCCGGA ATTCGGTCCG GAGCTTTATC GTTTGCCTTT GCCGGACAAG ATCCTGCGCA ATCATTGGTA CAAGTATGAA GTCGAGATT | 960 999 |
|----|---|---------------------------------------|
| 5 | (2) INFORMATION FOR SEQ ID NO:194 | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 945 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| | (ii) MOLECULE TYPE: DNA (genomic) | |
| 15 | (iii) HYPOTHETICAL: NO | |
| | (iv) ANTI-SENSE: NO | |
| 20 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| 25 | <pre>(ix) FEATURE: (A) NAHE/KEY: misc_feature (B) LOCATION 1945</pre> | |
| 20 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:194 | |
| 30 | ATGATCAGAA CGATACTTTC ACGATATGTA TCCTCGAACT TTTGGAGTCG GGGAGCTACC TTTTTTTTCA CGATTTTCCC GGCCTTCATC CTCGCCGCTA CTGCTTTGCC GGCTTGTGGA GGGGTACTG CTTCAGGCTC CGATCGTACG CTGGCTGTGA CCATCGAGCC ACAGAAATAC TTCATCGGAGT CATTGCGGA TAAGTCGGTG CAGTTGGTACC GGCCGCAGC AATCCGGAGG AATACGACCC TTCGCCTACC GTGATGAAGC GTTTGTCCGA AGCAGATGCC TACTTCTATA TAGGAGGACT GGGGTTCGAG CAAAGAAATC TCGCTGCCAT TCGGGACAAT | 60 120 180 240 300 360 |
| 35 | AACCCTAAGC TCCCTCTTTT CGAAATGGGC AAAGCCTTGG CGGATGCCGG AAGTGCAGAT CTCCACGGCT CCTGCACAGA TCATTCTCAT ACAGACCTGC ATGCCCATGA TCCGCACTAT TGGAGCAGTG TGGTAAGGGCACTA AGTCGTGCTG CATACGACGC GCTTGTGGAG CTTTATCCGA ACGAGAAAGA CAAATGGGAC AAAGGCCACA ACCGTCTCAA CGGACGTATC GACAGCGTGA AGAGACTCGT CGATACCATG TTTGCCAATG GCAAAGCAGA CAAAGCCTTC | 420 480 540 600 660 |
| 40 | GTCATATATC ACCCATCGCI CAGCTTTTTC GCCCAAGAGT TCGGCCTGCG GCAGATCGTC ATAGAGGAAG ATGGGAAAGA GCCTACGGCT GCCCACCTTC GTCGTGTGAT CGATCAGGCA CGTGCCGATG GTGTCAGAAT CGTATTTATC CAACCCGAAT TTGAAACGCG TCAGGCGGAG GACATCGCAC GCGAGATCGG TGCTCGTCCG GTAAGGATCA ATCCTCTGCG CAGCTCGTGG GAGGAGGAAA TTTTACATAT TGCTCGCGCT TTGGCTCATG AACGG | 720 780 840 900 945 |
| 45 | (2) INFORMATION FOR SEQ ID NO:195 | |
| 50 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2544 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 55 | (ii) MOLECULE TYPE: DNA (genomic) (iii) MYPOTHETICAL: NO | |
| | (iv) APTI-SENSE: NO | |
| 60 | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | |
| 65 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 12544 | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:195 | |
| 70 | ATGATOGGAA AAAAAATOTT TTTTATOOTG CTGGCGCTCA TTGCGTTCAG TGGGCTGAAC GCAGCGACAG ACACTGAGTT CAAGTACCCG ACCGATGCCA ATATCATCGG TCACGTCAAA GACAGCAAGA CGGGTGAACA CCTTGTCGGT ATCACTATTG CTATCAAAGG CACTACCTTT GGTACATCTA CAGATGCAAC CGGGCACTAC TATCTTCGTA ACTTGCGTCC GGGTGAGATC ACTTTGATTA TGCGTGGCAT GGGCTATAAG AGCCAGGAGC GCGTAGTCCG CGTAGAAAAG | 60 120 180 240 300 |
| 75 | GACAAGACTA TCGAGGTGAA TTTCGAAGCA GAAGAGGATG CCATCAATCT GGACGAAGTC GTGATTTCGG CCAACCGCGA ACTGACGCTT CGCCGTCTTG CTCCTACTCT GGTAAATGTA | 360 420 |

```
TTGAACGAAA AAGTCTTCTC GCAAGTCAAT GCTTCTAACC TGGCTCAAGG CTTGTCATTC
                                                                                          480
         CAGCCGGGAG TTCGTGTAGA GAACAACTGT CAGAACTGTG GTTTCAATCA AGTTCGTATC
                                                                                          540
        AATGGACTGG ATGGTCGTTA TGCACAGATC CTCATCGACA GCCGTCCCAT CATGAGTGCC
                                                                                          600
        CTTGCCGGTG TTTACGGTCT GGAGCAGATC CCTGCCAATA TGATCGAACG TGTGGAGGTA GTACGTGGTG GAGGATCGGC CTTGTACGGT TCTTCTGCTA TTGCCGGAGT GGTGAATATC
                                                                                          660
  5
                                                                                          720
        ATCACCAAGG AACCTTCTCA CAATTCTTTC ACATTCAATG AATCTCTGAG CTTTACCGGT
                                                                                          780
        TTCAGCAAGC TGGATAACAA CACGAACTTC AATGCCTCCA TCGTCAGCGA TGACAACCGT
                                                                                          840
        GCCGGTGCCA TGGTATTCGG GCAGGCTCGT TACCGCAACC ATTGGGATGC TAACAATGAC GGTTATTCCG AATTGGGTAA AATAGATGCC CGCTCGCTGG GAGCGCATTC TTATTTGCGC
                                                                                          900
                                                                                          960
 10
        TTGAGCGACT ACAGCAAATT GACGGGAGAG TTTCACACGA TCAGTGAATT CCGCCGTGGT GGCGATCGTA TCGATTTGCC TCCTCACGTA GTGGGTGTAG CTGAACAAAC TGACCATAGC
                                                                                         1020
                                                                                        1080
        GTATTTAGCG GAAACTTGAA ATACGATCTC TTCTCTTCCA ACTATAAACA CCACTTCCAG
                                                                                        1140
        GCTTATACTT CCGGACAGAT CGTAAATCGC AAGAGCTATT ACGGAGGTAT CGGAGAGATT
                                                                                        1200
        GACGTCAATG GCCACCCGG TGGTACGGAA GGCTACCCTA TCCCTCAAGA TCAATACGGC
                                                                                        1260
 15
        AATAATTATG GCGTGACCAA AGGCAAGACA TATATGGGCG GTATCCAGTA CAGCTACGAC
                                                                                        1320
        TTGGACAAAT TCCTCCTCAT GCCTTCGCAA CTTTTGTTCG GAGCCGAATA TACGCGTGAT
                                                                                        1380
        GAACTCAATG ACGTGATGCC CATCCTTTCA TGGCAGACCG GCGAGGATGC CAATGGGAAT
                                                                                        1440
        ACCATTCCCC TCTATCCCGA ATTGGATCAG AATATCAACA ACTACAGCCT ATTCGGTCAG
                                                                                        1500
        AACGAATGGA AAAATGACAG ATGGAGCATC CTTGTTGGCG CTCGCTTGGA CAAGCATAGC
GAAGTCAAGG ATATGATTCT GAGTCCTCGT ACCACACTGC GTTTCAACGT GAATCCGGAC
                                                                                        1560
 20
                                                                                        1620
        ATCAACCTGC GCGCTACATA TGCAAAAGGG TTCCGCGCAC CGCAGGTATT CGATGAAGAC
                                                                                        1680
        TTGCACGTAG GGGTTGTAGG CGGTGAGGCA CAGAAAGTAT TCAACGATCC GAACCTCAAG
                                                                                        1740
        CCTGAAATTT CTCATGCATT CAGTTTGAGT GCCGATATGT ATCATCGTTT CGGTAACGTC CAGACCAACT TCCTTGTGGA AGGCTTCTAT ACTCGTTTGC TGGATGTATT CACCAACGAG GAGCAGCCTG ATCAGCACGA TGGCATCAAA CGCTACACGC GTATCAACGG TAGCGGAGCC
                                                                                        1800
                                                                                        1860
25
                                                                                        1920
        AAAGTATTOG GTCTCAATCT GGAAGGTAAG GTCGCATACA AGTCCTTCCA GUTCCAAGCC
                                                                                        1980
        GGTCTTACCC TGGCCAGCAA CAAATACGAC GAAGCACAGG AGTGGGGTCT GAATACGGTG
                                                                                        2040
        AAAGACACCA ACGGAGCTTT TGTTACCGAG GCCAATGCAA ATGGACAACA GGAATACAAG
                                                                                        2100
        AACGAATCCA TGACGGATAC GCAGATCACC CGTACCCCCA GCGTATACGG TTATTTTACT
                                                                                        2160
30
        TTGGCCTACA ATCCTGCTCA CTCATGGAAC ATAGCCCTTA CGGGAGCATA TACCGGTCAG
                                                                                        2220
        ATGTATGTAC CCCACGCTAT CGAATATGGT GTGAAGTCTG CCGAACTGGA TATTATGCAG AACAATCCTG AGATTACCGA CGAAACCGGA AAGGCTCCCC GTATTGATGA GCTGAAGAAG
                                                                                        2280
                                                                                        2340
        ACACCTGCAT TCTTCGATTT GGGCTTGAAA GTGGGTTATG ACTTCCACGT ATTCCAGGCT
                                                                                        2400
        ACTGAGGTTC AACTCTATGT AGGTATGAAC AATATCTTCA ACTCTTTCCA GAAGGACTTC
                                                                                        2460
35
        GATCGTGGAG CTGCACGTGA CAGCGGATAT ATCTATGGTC CTACGCAGCC GCGTACAGGC
                                                                                        2520
        TACATGGGCT TGGTAGTGAA GTTC
                                                                                        2544
        (2) INFORMATION FOR SEQ ID NO:196
40
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 606 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
45
                    (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
50
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
55
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...606
60
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:196
       ATGACAGTAA AGCGCGCAGT GCGAATAGCA CTTCTCACGC TGATAGGCAT TCTTTTTCC
       TCACCTTCTC TTGTTCGGGC GCAAAGTCTT TTCAGCACCG AACATGTCTT GCAACTATAC
                                                                                        120
       AACAAGATAC TCTATGGAGA GTCGGCGGCG GATACCGTCG CAGAGAAAAC GGCAGGTGAG
65
                                                                                        180
       TOGGCATTTC CTTTTATAGA CAAACTCATC AATCTOGGCC GCACTTTCCT CGGCAAACCA
                                                                                         240
       TATCGCTATC GCGGTCCTTC CCCATGGCCG ATGGACTGCT CGGGCTATGT GTCTTACCTC
                                                                                         300
       TACTCCAAAT TCGACATCAA ACTCCCACGT GGTGCGGCAG CACAGAGCCA ATATACGAAT
                                                                                        360
       CCTATCGAGC GCGAGGATGT TCGTCCGGGC GACCTCCTTT TTTTCAAAGG CCGCAATGCA
                                                                                        420
       CGCAGCAACC GTATCGGGCA TGTAGCTTTG GTCGTATCTG TCGATGAAGA TGATATTACC
                                                                                        480
70
       ATGATGCACA GCCGCAATTC GCGAGGGATC GTGATCGAAA AACTCAATCG CAGTGCATAC
                                                                                        540
       TTCTCCCGTC GCTTGGTGAG CTATGGCAGG GTACCCGGAG CCAAGAGAGT GATCCCACGA
                                                                                        600
       AAAAGT
                                                                                        606
```

75

(2) INFORNATION FOR SEQ ID NO:197

```
(i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1365 base pairs
                  (B) TYPE: nucleic acid
 5
                  (C) STRAHDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
10
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
15
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1365
20
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:197
                                                                                    60
      ATGAAACGGA CAATCCTCCT GACGGCACTG ACCGTCCTAT CTTCGCTCTC CTTGCTTCGT
      GCACAAAATG AATCCGAAGC ATCAACCAAT CCGATGTCAG GCCTCTCCCT GGAAGACTGT
                                                                                   120
25
      ATCCGGATAG CCAAGGAGCG CAACCTGAAT CTGCGCAGAC AGGAGATCGA ACAAGAAAAC
                                                                                    180
      CGAATCATTA GTCTCGATGC AGCACGACAC AGTTTCCTGC CCTCGGTCAA TGCAGGCATC
                                                                                   240
      GGACACAACT ATAGCTTCGG ACGTTCGAAA GACAAAACGG GAGTAACCGT AGATCGCTCC
TCGATGAATA CCAATCTCAG CATCGGAGCT TCGGTGGAAG TATTCAGCGG CACACGTCGT
                                                                                    300
                                                                                    360
      CTGCACGACC TCAAGCAGCA AAAGTACAAC GTGGAGGATG GTATAGCCCG ACTTCAAAAA
                                                                                    420
30
      GCGCGTGAAG ACCTCAGCCT GCAAATCGCG GCTCTCTATA TCAATTTGCT CTTCCGTCAG
                                                                                    480
       GAAATGACTC GTACGGCAGA AACACAGTTG GCACTGATTC GCGAGCAACG CAATCGCACG
                                                                                    540
       GCCGAAATGG TTCGCGTAGG TAAATGGGCA GAGGGTAAGC TCCTCGACAT AAATGCCCAG
                                                                                    600
      ATGGCCAAGG ACGAACAACT TCTCGTACAA TATCGTTCGG AGGAGGAGCT GGCTCGTCTG
                                                                                    660
      GACTTGGGGC AAGCCCTCGA ACTGGAGCAC CCCGAAAGCA TTGCAGTCAA GGCTCCCGAC
                                                                                    720
35
      ACAGACGTTC TCGTAGCAGA AAGGTTGGGA TCTCTCCTTG CTCCCGAAGA GATCTATCGC
                                                                                   780
      ACGGCTCTCG GCTTGAAACC GGCACTGCAT TCGAGCGAGC TGCAAATAGC TTCGGCACGC GAAGGTCTGG CCTCGGCTCG TGCGGCATAC TTCCCGACGC TCAGCCTCTC TGCCGGATAC
                                                                                    840
                                                                                    900
      AGCAACGGTT ACTTCCGCGA CCTCGGCAAG GAGTATGCCG CCATCAACCC CTCCTTCTCC GAACAGTGGA AGAACAACGG CAGCTACAGT ATCGGACTCT CTTTGAATAT CCCCATCTTC
                                                                                    960
                                                                                  1020
40
                                                                                  1080
      TCTGCCATGC AAACGCAAGA TCGCGTTCGG AGCAGTCGCC TGCAAATACG CTCAAGCGAG
       CTTCGACTCG TCGAAGAGAA AAAAGCCCTC TATAAAGAGA TCAGGCAAGC ATACAGCAAT
                                                                                  1140
       GCCGTGGCAG CCGATAAGGC CATCGCAGCA GCCGAAAACA GCAAGGCCGC TACGCTCAAG
                                                                                  1200
       GCATACGAAT ACGCTCGCGA CAGCTTCGAG GCAGGGCGCT TGTCTGCCTA CGAATATGCC
                                                                                  1260
       GAGGCAAAAA CAAAATACGC CCTCAGCCAA GTGGAAGAAC TTCGTGCCAA GTATGACTTC
                                                                                  1320
45
      ATATACAAAG CCAAAGTTTT GGATTTCTAT CAGGGCAAAG ACTTC
                                                                                  1365
       (2) IMFORMATION FOR SEQ ID NO:198
50
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1332 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
55
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
60
           (iv) AUTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROHONAS GINGIVALIS
65
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1332
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:198
70
       ATGCGTTTCC AACATTATCT CATCTGTACG GCTGCCGTAG CGGCTTTGGC TGCGAATCCC
                                                                                     60
       CTTACGGGCC AATCGAATAT GACCCTCGAA GAGTGCATAG ACTATGCACG CCGGCACAGT
                                                                                    120
       TCGGCCGTGG CGCTGTCCGC TGCGGAACTG GAGCAGTCCA AGGCCGATTA CCTTCAGGCC
                                                                                    180
       GTCGGCAATT TTCTGCCCCG TGTATCGGCC GGAACCGGTG CTTCGTGGAA TTTCGGACGC
                                                                                    240
75
       GGATTGGATG CCGAGACGAA TACCTACACC GACATCAACA GCTTCAACAA TTCGTACAGC
                                                                                    300
```

```
ATACATGCCA CGATGACCCT TTTCGACGGT TTGCAGAGTG TCTATCGACT GCGGATGGCG
       CATGCACGCC GGGAGGCTTC GCGCCTCTCC GTTCGCGAGC AGCAGGAGCT GGCAGCTCTC
                                                                               420
       GGCACCACGG AGGCCTACTA CGACCTCGTC TATGCGCGCC AAATGCAAGA GCTGGCCATG
                                                                                480
       CAGAAGTACG AGGAGAGCAG CCGCCTCCAC CGGCAGACGG CTCGAATGGA AGAGCTGGGG
                                                                               540
 5
       ATGAAGAGTC GTCCCGATGT CCTCGAGATG CAGTCGCGAA TGGCCGGTGA CCGTTTGGCC
                                                                               600
       CTGACTCAAG CGGACAATCA GTGCATCATC GCTCTGATCC GGCTCAAAGA AAAAATGAAC
                                                                               660
       TTCCCCATCG ATGACGAACT CGTCGTAGAC GATATGCCGG CTGACAGTCT CTCCGCCGAC
                                                                               720
      ATGGCCGAAT CGGACAGCTC GGCCGGCGTC TTCGCCCGTG CTGCCCATCA TCATCCCGTC
                                                                               780
       CTCCTCCGTG CCAAACTCGA CGAGCAGGCT GCCACCGACC GTTTGCGAGC CGCGCGAGGT
                                                                               840
10
       GCATTCCTGC CGAGTGTGTC GGTATCCGGA GGATGGAACA CGGGATTCTC ACGCTTTTTG
                                                                               900
       AATGGATCGG ACTATACGCC CTTCAGCGAG CAGTTTCGGA ACCGTCGGGG GGAATACGTC
                                                                               960
      AGTCTGAATC TGAGTATCCC CATCTTTTCG GGATTCAGCC TTGTGAGCCA TCTGCGTCAG
                                                                              1020
      GCGCGTGCCG AACGCAGGGC GGCAATCGTC CGACGGGGCG AAGCGGAGCG CAGGCTCTAC
                                                                              1080
      AGCGAGATCG CCCAAGCCAT GGCCGACCGG GATGCCGCTC TGGCTTCCTA CCGCCAGGCG
                                                                              1140
      AAGGAGCATA CCGACGCCAT GCAAACCGCT TACGAAGCCG TCTTGCAGCG TTATGAGGAG
GGGCTGAATA CGGCCATCGA CCTGACCACT CAGGCCAATC GGCTCCTGGA TGCCCGTGTG
15
                                                                              1200
                                                                              1260
      CAGCGACTGA GAGCGGCCAT GACCTACCGG CTCAAATGCA AACTCATAGC CTATTACGGC
                                                                              1320
      TGCCTTTCGG AC
                                                                              1332
20
       (2) INFORMATION FOR SEQ ID NO:199
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2820 base pairs
25
                 (B) TYPE: nucleic acid
                    STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
30
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
35
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
40
                 (B) LOCATION 1...2820
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:199
      ATGAACAAAT TTTACAAATC ACTTTTGCAG TCAGGACTGG CTGCCTTCGT GTCGATGGCA
45
      ACTGCACTGA CCGCTTCTGC ACAGATTTCG TTCGGAGGGG AACCCTTGAG TTTCTCTCA
                                                                               120
      AGATCCGCCG GAACGCATTC ATTCGACGAT GCAATGACTA TCCGCCTTAC TCCGGATTTC
                                                                               180
      AATCCGGAAG ACCTGATCGC ACAGAGCCGT TGGCAATCGC AAAGAGATGG CCGGCCCGTC
                                                                               240
      CGGATAGGAC AAGTAATACC GGTGGATGTG GACTTTGCAT CCAAGGCTTC GCACATCTCT
                                                                               300
      TCCATCGGAG ACGTAGATGT ATATCGCCTG CAATTCAAGT TGGAAGGAGC CAAAGCCATT
                                                                               360
50
      ACGCTTTATT ACGATGCATT CAATATTCCG GAGGGCGGAC GCCTCTATAT CTATACCCCC
                                                                               420
      GACCATGAAA TIGIGITGGG AGCATATACG AACGCCACTC AICGCCGCAA CGGAGCTITI
                                                                               480
      GCCACAGASC CGGTACCGGG GAGTGAGCTT ATTATGGATT ATGAAGTGTC TCGCGGAGGG
                                                                               540
      ACTTTGCCTG ACATCAAGAT CTCCGGTGCG GGTTATATAT TCGACAAAGT CGGCGGACGC
                                                                               600
      CCCGTAACGG ATAACCATTA CGGGATCGGT GAGGACGATT CCGATTCGGA TTGCGAGATC
                                                                               660
55
      AACATCAATT GTCCTGAAGG TGCAGACTGG CAGGCAGAGA AGAACGGTGT GGTGCAAATG
                                                                               720
      ATCATGGTAA AAGGACAGTA TATCTCAATG TGCTCAGGCA ACCTGCTCAA TAATACGAAA
                                                                               780
      GGAGACTTTA CTCCGCTGAT CATTTCTGCC GGACACTGTG CTTCCATAAC AACCAATTTC
                                                                               840
      GGTGTAACGC AATCCGAGTT GGATAAGTGG ATCTTCACTT TCCACTATGA AAAAAGAGGA
                                                                               900
      TGCAGCAATG GTACATTGGC CATCTTCCGT GGCAACAGTA TCATCGGAGC TTCCATGAAG
                                                                               960
60
      GCTTTCCTCC CGATCAAAGG TAAATCCGAT GGTCTCTTGC TGCAACTCAA CGATGAAGTC
                                                                              1020
      CCTCTGCGCT ATCGTGTCTA TTACAATGGA TGGGACAGTA CGCCCGATAT TCCCTCGAGC
                                                                              1080
      GGTGCCGGTA TTCATCATCC GGCCGGAGAT GCCATGAAGA TTTCCATCCT AAAGAAGACT
                                                                              1140
      CCGGCTCTGA ATACATGGAT CTCCTCCAGT GGTTCCGGAG GGACTGACGA TCACTTCTAT
                                                                              1200
      TTCAAATAGG ATCAAGGTGG TACGGAAGGA GGATCGTCCG GTTCTTCTCT CTTCAATCAG
                                                                              1260
65
      AATAAGCACG TGGTCGGCAC ACTGACCGGA GGTGCCGGCA ATTGTGGCGG GACGGAGTTC
                                                                              1320
      TACGGCAGAC TGAACAGTCA TTGGAACGAG TATGCATCCG ATGGCAATAC GAGCCGCATG
                                                                              1380
      GACATCTATC TGGATCCCCA AAACAATGGC CAGACGACCA TCCTCAACGG AACGTATCGT
                                                                              1440
      GACGGTTATA AGCCTTTGCC CTCTGTGCCC CGGCTATTGT TGCAGTCTAC AGGCGATCAG
                                                                              1500
      GTCGAATTGA ATTGGACGGC TGTTCCTGCC GATCAATATC CATCATCTTA TCAGGTCGAA
                                                                              1560
70
      TACCACATAT TCCGAAATGG AAAGGAAATA GCTACGACAA AGGAGTTGTC CTATTCGGAT
                                                                              1620
      GCCATCGACG AAAGTATTAT CGGTAGCGGT ATCATTCGAT ACGAAGTAAG CGCACGCTTC
                                                                              1680
      ATTTATCCCT CGCCGTTGGA TGGAGTGGAA TCTTATAAGG ATACGGACAA GACTTCTGCC
                                                                              1740
      GACCTTGCCA TAGGAGACAT TCAGACCAAG CTGAAGCCGG ACGTAACACC TCTCCCCGGA
                                                                              1800
      GGAGGAGTAT CATTAAGCTG GAAAGTTCCT TTCTTAAGCC AGTTGGTTTC CCGATTCGGA
                                                                              1860
75
      GAAAGCCCCA ATCCTGTGTT CAAAACCTTT GAAGTGCCCT ATGTTTCTGC CGCAGCCGCA
                                                                              1920
```

```
CAAACCCCCA ATCCTCCGT TGGCGTAGTC ATTGCAGACA AGTTTATGGC CGGTACATAT
                                                                                   1980
       CCCGAAAAGG CTGCTATCGC TGCCGTTTAT GTAATGCCAT CCGCTCCGGA CTCTACTTTC
                                                                                    2040
       CACCTCTTCC TCAAGAGCAA CACAAACAGA AGATTGCAGA AGGTGACAAC TCCCTCCGAT
                                                                                    2100
       TGGCAGGCCG GAACATGGTT GAGGATCAAT TTGGATAAGC CGTTCCCGGT GAATAATGAC
                                                                                    2160
 5
       CATATGCTTT TTGCCGGTAT CAGAATGCCT AATAAGTACA AGCTCAATCG TGCTATCCGT
                                                                                    2220
       TATGTAAGAA ATCCGGATAA CCTTTTCTCC ATTACCGGTA AGAAGATTTC ATATAACAAC
                                                                                    2280
       GGAGTCTCTT TCGAAGGCTA CGGAATACCC TCGCTCTTGG GCTATATGGC TATCAAATAT
CTGGTGGTAA ATACCGATGC TCCGAAGATC GATATGTCGC TTGTACAGGA GCCTTATGCT
                                                                                    2340
                                                                                    2400
       AAGGGAACGA ATGTGGCTCC ATTCCCCGAA TTGGTCGGCA TATATGTCTA TAAGAACGGA
                                                                                    2460
10
       ACATTTATCG GCACACAGGA TCCATCCGTC ACAACTTATT CGGTTTCAGA CGGAACAGAG
                                                                                   2520
       AGCGATGAAT ACGAAATAAA ACTGGTATAT AAGGGATCGG GCATTTCGAA TGGCGTTGCT
                                                                                   2580
       CAGATTGAGA ATAACAATGC TGTCGTTGCA TATCCGTCTG TTGTAACAGA TCGTTTCAGC
                                                                                   2640
       ATTAAGAACG CTCATATGGT TCACGCTGCC GCCCTCTACT CATTGGATGG CAAGCAGGTT
                                                                                   2700
       CGTTCTTGGA ACAACCTCCG CAATGGCGTG ACATTCAGTG TTCAAGGACT TACGGCCGGT
                                                                                   2760
15
       ACTTATATGC TCGTTATGCA GACGGCAAAC GGCCCTGTGA GCCAAAAGAT CGTGAAGCAG
       (2) INFORMATION FOR SEQ ID NO:200
20
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2010 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDMESS: double
                  (D) TOPOLOGY: circular
25
           (ii) NOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
30
           (iv) AUTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
35
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...2010
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:200
40
      ATGAAATATC TTATCAGACT CTTCTTATCA TTGATGTTAC TCTCTCTG GACGGGCTGT
      ACACACGAGG AGCTCTCTAT TTGCGATGGC GAGAATACGC TTGTTTTACG CGTAGAGACC
                                                                                    120
      GGTAAAGCCC CAAATGCTCG TGCCACAGAA CCCGGTCAGG GCATATACAA TGAGAATAAA
                                                                                    180
       GTAGGCTCCA TTTCTGTGCT CTTCTATTTA GAGGGACAAC TTCGTTGGCA GGTGAAGTCT
                                                                                    240
45
       ACAGACTATC AAATCCATGA AGGGGCCTAT ATCATTCCGG TCAAAGAGCA AATGCGACCA
                                                                                     300
       CTATTCAATG GCAACAACAA CTTCAGCATC TATGTAGTGG CCAATCTCGA TTTCAATGCT
                                                                                    360
      CCGGCCACAG AAGCTGCGCT TTCTCAATTT GTGGTAGAGA AATCTATTGA AGTCTCTTCT
                                                                                    420
      ACGACAGCCC CTGCCGATTT CGTAATGCTT GCTCATGGCA ATAAGCAGAT CAATATGGCT ACGACAGAAG GGAAACTGTT GGGGGATTAT AAACTCAAAC GAGTGGCAGC AAAGATTCGC
                                                                                    480
                                                                                    540
50
      ATGATAAAAC CCACCATCAA TGTGCAAGGA TATGAAGTGG TCGGAAATAT ACAGGCAAAG
                                                                                    600
      TTTCGCAATT CGGTAACGAA GGGGTTCCTT ACCACAGAAG CTCAAGAGAT CCCAGCTGCT
                                                                                    660
      GCATCCTATA AGACATCGGA ATATCTTGAT ATTGCAGAGT CGGCACCTGC CAATTCTATC
                                                                                    720
      CATTTCTATT CTTACTATAA CAAATGGACA CTCTCCACAC CGGAGAAGCG ACCGGAATTC
                                                                                    780
      TTCATCATGG TCAAATTCAA AAAGACAGGA CAGCCGGACA ACACAGCCAA ACCGTACTAC
                                                                                    840
55
      TACAGAGTGC CCCTCGAATC TCAGGACAAT CAGGTCAAGA GCAATGTCCT CTATAATCTG
                                                                                    900
      AATGTGAAAA TCGAAATCTT GGGTTCTTTA CAAGAGCCGG AAGCTGTTTC TGTAAACGGC
                                                                                    960
      ACACTCGCAA TAGAAGAATG GATTCTCCAT CAGGATGCAT TCAATCTGCC TGCCACCAAT TACTTGATAG TGGAACAGCA CGAAATCTTC ATGAATAACG TGAACACATA CTCGGTGAAA
                                                                                   1020
                                                                                   1080
      TATCAAACTT CGCAGAAACC AATCAGCATT AGCATACAGT CAGTTACCTT TAGCTACGTC
                                                                                   1140
60
      TCTTCTGATG GCACTCAGCA CAATGATCTT GTAGCAAGTA GTAGCGACCA GTATCCTACG
                                                                                   1200
      ATTACAAGCG ATAATACAAG CATCATAATC ACTTCCAAGA TACCGGTTAA TAACGTACCA
AAGAAGATCG TTTTTGAGGT AACTAATGGG GTAGCCGGTT TGAAAGAGAC TGTCACAGTA
                                                                                   1260
                                                                                   1320
      CTCCAATATC CTGCACAATT TATTGTCAAT ACACTTGGCA CAGCATCGGC ATGGAGACCA GACGGATCTT TGGCTCCGGG GCTTAACAAT AAAGCGATTT ACCATGTCGT AGTACTGGTT
                                                                                   1380
                                                                                   1440
65
      CCACCCGAGA ATTTATTTGA AGATGGGACA CAGACAATCA TCGGTTATCC CCCCACTGAA
                                                                                   1500
      ACAATTTCTT TTCATAAGAA AGAGAACAAT ACCTATCCGA TAGTATGGTC TGACACAAAT
                                                                                   1560
      ACGACAAAAC AGGACCTTGA GACATCAAGA ATGATTTCAC CTTCCTTTGA GTTAGCCTCC
                                                                                   1620
      CAACTTGGGG CTACTCTCCC GATGCCCTAT CTCGAGTATT GGCCAGGGAC ATCATATCTC
                                                                                   1680
      CTTGACTATT CGGGAAACTA TAATAATAAG AGATACGCCT TGTTTAATTG CGCTTTTTAC
                                                                                   1740
70
      TGGGAGAAAA GAAAAGTTAA TAACGAAGAA ATTAAATTCG ATGACTGGCG TTTGCCGACA
                                                                                   1800
      GAAGCTGAGA TCAAATTGAT AGATAAGCTG CAACATAATG AGCAGAGTGC TGTCCAAGCT
                                                                                   1860
      ATCATGACAG GGAATTATTA TTGGGATAGT TACTCTGCAA ATGGGTCTTA TAAAATGCAA
                                                                                   1920
      GGAGGAGGG GCCAAGGAAA TTCCTCCAAA GCCTATGTTC GTTGCGTGCG GGATGTGAAA
                                                                                   1980
      AAGCCGATTC GTGACAAGAA GTCAGGTAAG
                                                                                   2010
75
```

```
(2) INFORMATION FOR SEQ ID NO: 201
            (i) SEQUENCE CHARACTERISTICS:
 5
                  (A) LENGTH: 3846 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
10
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
15
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (im) FEATURE:
20
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...3846
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:201
25
      ATGCGAAAAA TTTTGAGCTT TTTGATGATG TGCTCTCTGC ATTTAGGTCT ACAATCTCAG
      ACTTGGCATG GAGATCCGGA CTCAGTGGCA GCCCTACCTT CTATCGGTAT TCAAGAGTCA
                                                                                   120
      AGTTGTACCC GAATCACGTT CGAGGTTGTT TTCCCCGGAT TTTATAGTGT GGAAAAACGA
      GAAGGCAACC AAGTCTTTCA GCGCATTTCC ATGCCGGGTT GTGGCTCGTT TGGGAATCTG
      GGCGAAGCTG AATTGCCTGT TTTGAAAAAG ATGATAGCCG TTCCGGAATT TTCAACAGCT
                                                                                   300
30
      AACGTTGCTG TAAAAATCAA AGAGACGGAG ACATTCGACA ATTATAATAT CTATCCTAAT
                                                                                   360
       CCTACCTATG TCGTAGAGGA GTTGCCTGAG GGGGGGACTT ATCTGGTAGA GGCTTTCGCG
                                                                                   420
      ATAAACAATG ACTATTATAG CCAAAATGTA AGCCTCCCTT CTACTCACTA TGTCTATTCT
CAAGACGGGT ATTTTCGCTC ACAAAGATTT ATCGAAGTTA CCCTGTATCC TTTTCGATAC
                                                                                   480
                                                                                   540
      AACCCTGTCC GACAAGAAAT TCTATTTGCA AAAAAAATCG AGGTTACAAT AACTTTCGAT
                                                                                   600
35
      AATCCTCAGC CACCTTTACA AAAAAACACC GGCATATTTA ACAAAGTAGC CTCCTCTGCA
                                                                                   660
      TTTATTAATT ATGAAGCTGA TGGCAAATCG GCGATAGAAA ATGATATGGT GTTCAGTCGT
                                                                                   720
      GGTACAACAA CGTACATAAG CGGAAATGTT GCCAGCAACC TCCCTCAGAA CTGTGACTAC
                                                                                   780
      TTGGTTATTT ACGATGATAT GTTCAACGTA AATCAACAAC CACACGACGA AATCAAACGG
CTGTGCGAAC ATAGAGCCTT CTACAACGGC TTTGATGTAG CTGCTGTAAG TATAAAGGAC
                                                                                   840
                                                                                   900
40
      GTATTGAATA GCTTCCCATC AAATGCCACC TCATACATCA ACGAAACTAA ACTGAAAAAT
                                                                                   960
      TTCATTCGCT CAGTTTACAA CCAAAGCAAT GCGAAGAGGA CTTTAGATGG CAAACTGGGA
                                                                                  1020
      TACGTGCTAC TGATCGGAAA ACCATTGAGC AAATATTTGG CTGACACTGA TAATACAAAA
                                                                                  1080
      GTCCCAACCT CTTTTATTCA TAATGTCTCC TTAATTCCAA GTCATCCAAC TTTTGGTTCC
                                                                                  1140
      ATATGCGCCT CCGACTATTT TTTTAGTTGT GTTTCGCCCC TTGATACTGT CGGCGATTTG
                                                                                  1200
45
      TTTATCGGTC GATTTAGCGT CACCAATGCT CATGAATTGC ACAATCTGAT TGAAAAGACT
                                                                                  1260
      ATCAACAAAG AAATCTCATA TAATCCTATT GCACACAAAA ATATTCTTTA CGCAGAAGGG
AAAGGCTGCG ATGCTCCAAT CTTACGTTTA TTCTTAAAAG AAATCGCCTC TGGTTACACA
                                                                                  1320
                                                                                  1380
      GTCAACTCTA TCTTAAAATC TAATCAGGTC TCTGCAATAG ACTCGATATT TGACTGCTTG
                                                                                  1440
      AATAATGGTT CCCATCATTT TTATTTTAAC ACTCATGGAA TGCCGACTGT TTGGGGGATA
                                                                                  1500
50
      GGGCAGGGAC TCGACGTCAA TACTCTAACA GCCCGATTGA ACAATACATC TTCGCAGGGA
      TTATGTACGA GTCTATCATG TAGTTCGGCT GTAGCAGATT CAACTATTAG ATCGCTTGGA
                                                                                  1620
      GAAGTCCTGA CCACATACGC ACCTAACAAG GGATTCTCGG CTTTCTTAGG AGGAAGCAGA
                                                                                  1680
      GCCACCCAAT ATGCCGTTTA TTTAGAAGGC CCCTGTCCTC CGTCAGAATT TTATGAATAT
      TTACCTTATT CTTTATATCA CAATCTCTCG ACTGTTGTTG GCGAAATGTT GCTATCATCC
                                                                                 1800
55
      ATTATCAATA CTAATTCTGT TGATACGTAT TCGAAATTCA ACTTCAATTT GCTTGGCGAC
                                                                                  1860
      CCTGCACTAA ACATTATGGC TCATGGCATG GAGGTTAGTA ATTGTATTAC ACTACCAAAC
                                                                                 1920
      AACACCATTA TAAGCAGTCC GATAACAATA AAAAATGGTG GCTGCCTAAA AATACCGGAA
                                                                                  1980
      AAAGGAGTTT TGCATTTTAC TAATAATGGC TCCATACAAG TCATGTCCGG AGGAACTCTG
                                                                                  2040
      GAAATAGGCA ATCAGGCTAA AATATCCGGA GAGACCGGTG CTAACCCCAC CTTTATTACC
                                                                                  2100
60
      GTTTACGGCG ATGGTCTTGC GATTAACAAG CAGGTAGAGA TAGACAATAT AGACCGACTT
                                                                                  2160
      AACTTGTTTT CTACGCATTC GGTCATGCCC AAATTTCATT TTGACAGTGT GAAATTCAAC
                                                                                  2220
      AGTGCCCCGC TGTATACAAC GAACTGTATT GTGGAGATAA GCAATTGCGA ATTTACCAAT
                                                                                  2280
      CGAAGTGACA TTATTTCAAA GAATTGTGAC CTAAGCGTTG AAAACAGTAT GTTTAGCAGT
                                                                                  2340
      TCGGGGATAA CGGTATTCAA GCCTATGGCT ACAAGCTCCA TCACCGGATT ATCTACAAAA
                                                                                  2400
65
      GCAAAGATTA CCGACAATAC TTTTTTTGCG ACAGGAAACT TCGCCTACCA TATCACAAAC
                                                                                  2460
      ACGCCAGGCT TAACAGCAAC CTCCAATGCT GCCATCAAGT TAGACAATAT TCCTGAGTAT TACATTTCCG GTAATAAAAT AGTCAATTGC GATGAGGCTC TTGTACTAAA TAATAGTGGC
                                                                                  2520
                                                                                  2580
      AACAGAACGA ACAGACTCCA CAATATCACA CGGAATGTGA TAAAAAACTG TAGGATTGGG
                                                                                  2640
      AGCACGCTTT ATAATTCCTA TGGTATTTAC AACCGAAATA AGATCAGTAA CAATCATATA
                                                                                  2700
70
      GGAGTACGTC TCCTCAACAA CAGTTGTTTT TATTTCGATA ATGCTCCTGT AATCAATGAA
                                                                                  2760
      GAAGATAAGC AGACGTTTAT TTCTAATAGG ACTTGGCAGC TCTATTCATC AAACGGTACA
      TTCCCTCTCA ACTTCCATTA CAACAGCTTG CAGGGGGGAG ATACAGATAC ATGGATTTAC
                                                                                  2880
      AACGACACGT ATACGANTCG CTATATTGAC GTTTCAAATA ATCACTGGGG CAACAATGAT
                                                                                  2940
      TTGTTTGATC CGAATCAGGT TTTCAATACG CCAGACTTGT TCATTTGGAT ACCTTTTTGG
                                                                                  3000
75
      GATGGATTGC CAAATGGGAG ATCGGGCAAT AGCTCTGCTG AAGCAGTAGA ATTCCAAACA
                                                                                  3060
```

WO 99/29870 PCT/AU98/01023

```
GCATTGGACT GTATTGGCAA TAGCGATTAT CTTTCGGCAA AAGTGGCTCT CAAGATGATG
                                                                                           3120
       GTTGAAACCT ACCCGGAATC CGACTTTGCA ATAGCTGCTT TGAAGGAATT GTTCAGGATA
                                                                                           3180
       GAGAAAATGT CAGGCAACGA TTACGAAGGC TTGAAAGATT ATTTCAGATC CAATCCAACC
                                                                                           3240
       ATCATCTCTT CCCAGACTT TTGCAAGGC TTGAAAGATT ATTTCAGATC CAATCCAACC
ATCATCTCTT CCCAGACTT GTTCCCGACA GCTGATTTCC TGCGCGG ATGCGATATT
GTGTGTGAAA ACTATCAGTC TGCCATCGAT TGGTACGAAA ATCGCTTGAA TAGTGAAATC
TCCTATCAGG ACAGTGTTT TGCAGTCATT GACCTTGGTG ACATTTATTG GAATATCCAG
TTAGACTCAC TCAGAGGGAC TGGTATAGAT TTGAACATAC TTTCCTGTGA ACAAAGGAAA
TCGCTCGAAA GCCATCAAAA TGTAAAAAAT TATTTGTTGT CAACTCTTCC CGAATCAACA
GGTACTCTCC TGCCTCCATT AGAATGCAAC AAATCAAGGC TTGATAAATC CAAGATAAATC
TCGATTTTCGC CCAATCCAGC GAAAGCACTT CTAACACTTC CCAATCAACC
                                                                                           3300
 5
                                                                                           3360
                                                                                           3420
                                                                                           3480
                                                                                           3540
                                                                                           3600
10
       TCTATTTCGC CCAATCCGGC GAAAGCTGTT GTAACAATAA TCTACTATAC CGATAACCCT
                                                                                           3660
       TCCTGTTCTG TAATAAAAAT ATATGGAATA AATGGAGCCT CGGCTGATAT AACCGGGTTG
                                                                                           3720
       CCCAAACATC TATCCGAAGG TTATTACAGC ATACAGTTCA ATACATCCAA CTTTGATCCC
                                                                                           3780
       GGTTTCTACC TGGTAACGCT AAATGTTGAT CAGAAAATTA TAGATACGGA AAAATTACGA
       ATCAAA
15
       (2) INFORMATION FOR SEQ ID NO: 202
             (i) SEQUENCE CHARACTERISTICS:
20
                    (A) LENGTH: 3822 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
                    (D) TOPOLOGY: circular
25
            (ii) HOLECULE TYPE: DNA (genomic)
           (iii) HYPOTRETICAL: NO
            (iv) ANTI-SENSE: NO
30
            (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
35
                   (A) NAME/KEY: misc feature
                    (B) LOCATION 1...3822
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:202
40
       ATGATGTGCT CTCTGCATTT AGGTCTACAA TCTCAGACTT GGCATGGAGA TCCGGACTCA
                                                                                             60
       GTGGCAGCCC TACCTTCTAT CGGTATTCAA GAGTCAAGTT GTACCCGAAT CACGTTCGAG
                                                                                            120
       GTTGTTTTCC CCGGATTTTA TAGTGTGGAA AAACGAGAAG GCAACCAAGT CTTTCAGCGC
                                                                                            180
       ATTTCCATGC CGGGTTGTGG CTCGTTTGGG AATCTGGGCG AAGCTGAATT GCCTGTTTTG
                                                                                            240
       AAAAAGATGA TAGCCGTTCC GGAATTTTCA ACAGCTAACG TTGCTGTAAA AATCAAAGAG
                                                                                            300
45
       ACGGAGACAT TCGACAATTA TAATATCTAT CCTAATCCTA CCTATGTCGT AGAGGAGTTG
                                                                                            360
       CCTGAGGGGG GGACTTATCT GGTAGAGGCT TTCGCGATAA ACAATGACTA TTATAGCCAA
                                                                                            420
       AATGTAAGCC TCCCTTCTAC TCACTATGTC TATTCTCAAG ACGGGTATTT TCGCTCACAA
                                                                                            480
       AGATTTATCG AAGTTACCCT GTATCCTTTT CGATACAACC CTGTCCGACA AGAAATTCTA
                                                                                            540
       TTTGCAAAAA AAATCGAGGT TACAATAACT TTCGATAATC CTCAGCCACC TTTACAAAAA
                                                                                            600
50
       AACACCGGCA TATTTAACAA AGTAGCCTCC TCTGCATTTA TTAATTATGA AGCTGATGGC
                                                                                            660
       AAATCGGCGA TAGAAAATGA TATGGTGTTC AGTCGTGGTA CAACAACGTA CATAAGCGGA
AATGTTGCCA GCAACCTCCC TCAGAACTGT GACTACTTGG TTATTTACGA TGATATGTTC
                                                                                            720
                                                                                            780
       AACGTAAATC AACAACCACA CGACGAAATC AAACGGCTGT GCGAACATAG AGCCTTCTAC
                                                                                            840
       AACGGCTTTG ATGTAGCTGC TGTAAGTATA AAGGACGTAT TGAATAGCTT CCCATCAAAT
                                                                                            900
55
       GCCACCTCAT ACATCAACGA AACTAAACTG AAAAATTTCA TTCGCTCAGT TTACAACCAA
                                                                                            960
       AGCAATGCGA AGAGGACTTT AGATGGCAAA CTGGGATACG TGCTACTGAT CGGAAAACCA
                                                                                           1020
       TTGAGCAAAT ATTTGGCTGA CACTGATAAT ACAAAAGTCC CAACCTCTTT TATTCATAAT
                                                                                           1080
       GTCTCCTTAA TTCCAAGTCA TCCAACTTTT GGTTCCATAT GCGCCTCCGA CTATTTTTTT
                                                                                           1140
       AGTTGTGTTT CGCCCCTTGA TACTGTCGGC GATTTGTTTA TCGGTCGATT TAGCGTCACC
                                                                                           1200
60
       AATGCTCATG AATTGCACAA TCTGATTGAA AAGACTATCA ACAAAGAAAT CTCATATAAT
                                                                                           1260
       CCTATTGCAC ACAAAATAT TCTTTACGCA GAAGGGAAAG GCTGCGATGC TCCAATCTTA
                                                                                           1320
       CGTTTATTCT TAAAAGAAAT CGCCTCTGGT TACACAGTCA ACTCTATCTT AAAATCTAAT CAGGTCTCTG CAATAGACTC GATATTTGAC TGCTTGAATA ATGGTTCCCA TCATTTTAT
                                                                                           1380
       TTTAACACTC ATGGAATGCC GACTGTTTGG GGGATAGGGC AGGGACTCGA CGTCAATACT
                                                                                           1500
65
       CTAACAGCCC GATTGAACAA TACATCTTCG CAGGGATTAT GTACGAGTCT ATCATGTAGT
                                                                                           1560
       TOGGCTGTAG CAGATTCAAC TATTAGATCG CTTGGAGAAG TOOTGACCAC ATACGCACCT
                                                                                           1620
       AACAAGGGAT TCTCGGCTTT CTTAGGAGGA AGCAGAGCCA CCCAATATGC CGTTTATTTA
                                                                                           1680
       GAAGGCCCCT GTCCTCCGTC AGAATTTTAT GAATATTTAC CTTATTCTTT ATATCACAAT
CTCTCGACTG TTGTTGGCGA AATGTTGCTA TCATCCATTA TCAATACTAA TTCTGTTGAT
                                                                                           1740
                                                                                           1800
70
       ACGTATTCGA AATTCAACTT CAATTTGCTT GGCGACCCTG CACTAAACAT TATGGCTCAT
                                                                                           1860
       GGCATGGAGG TTAGTAATTG TATTACACTA CCAAACAACA CCATTATAAG CAGTCCGATA
                                                                                           1920
       ACAATAAAAA ATGGTGGCTG CCTAAAAATA CCGGAAAAAG GAGTTTTGCA TTTTACTAAT
                                                                                           1980
       AATGGCTCCA TACAAGTCAT GTCCGGAGGA ACTCTGGAAA TAGGCAATCA GGCTAAAATA
                                                                                           2040
       TCCGGAGAGA CCGGTGCTAA CCCCACCTTT ATTACCGTTT ACGGCGATGG TCTTGCGATT
                                                                                           2100
75
       AACAAGCAGG TAGAGATAGA CAATATAGAC CGACTTAACT TGTTTCTAC GCATTCGGTC
                                                                                           2160
```

PCT/AU98/01023

| - | ATGCCCAAAT TTCATTTTGA CAGTGTGAAA TTCAACAGTG CCCCGCTGTA TACAACGAAC TGTATTGTGG AGATAAGCAA TTGCGAATTT ACCAATCGAA GTGACATTAT TTCAAAGAAT TGTGACCTAA GCGTTGAAAA CAGTATGTTT AGCAGTTCGG GGATAACGGT ATTCAAGCCT ATGCTACAA GCTCCATCAC CGGATTATCT ACAAAAGCAA AGATTACCGA CAATACTTTT | 2220 2280 2340 2400 |
|------------|---|--------------------------------------|
| 5 | TTTGGGACAG GAAACTTGGC CTACCATATC ACAAACACGC CAGGCTTAAC AGCAACCTCC AATGCTGCCA TCAAGTTAGA CAATATTCCT GAGTATTACA TTTCCGGTAA TAAAATAGTC AATTGCGATG AGGCTCTTGT ACTAAATAAT AGTGGCAACA GAACGAACAG ACTCCACAAT ATCACACGGA ATGTGATAAA AAACTGTAGG ATTGGGAGCA CGCTTTATAA TTCCTATGGT ATTTACAACC GAAATAAGAT CAGTAACAAT CATATAGGAG TACGTCTCCT CAACAACAGT | 2460 2520 2580 2640 2700 |
| 10 | TGTTTTATT TCGATAATC TCCTGTAATC AATGAAGAAG ATAAGCAGAC GTTTATTTCT AATAGGACTT GGCAGCTCTA TTCATCAAAC GGTACATTCC CTCTCAACTT CCATTACAAC AGCTTGCAGG GGGAGATAC AGATACATG ATTTACAACG ACACGTATAC GAATCGCTAT ATTGACGTTT CAAATAATCA CTGGGGCAAC AATGATTTGT TTGATCCGAA TCAGGTTTTC AATACGCCAG ACTTGTTCAT TTGGATACCT TTTTGGGATG GATTGCCAAA TGGGAGATCG | 2760 2820 2880 2940 3000 |
| 15 | GGCAATAGCT CTGCTGAAGC AGTAGAATTC CAAACAGCAT TGGACTGTAT TGGCAATAGC GATTATCTTT CGGCAAAAGT GGCTCTCAAG ATGATGGTTG AAACCTACCC GGAATCCGAC TTTGCAATAG CTGCTTTGAA GGAATTGTTC AGGATAGAGA AAATGTCAGG CAACGATTAC GAAGGCTTGA AAGATTATTT CAGATCCAAT CCAACCATCA TCTCTTCCCA GAACTTGTTC | 3060 3120 3180 3240 |
| 20 | CCGACAGCTG ATTTCCTGTC TGCGCGATGC GATATTGTGT GTGAAAACTA TCAGTCTGCC ATCGATTGGT ACGAAAATCG CTTGAATAGT GAAATCTCCT ATCAGGACAG TGTTTTTTGCA GTCATTGACC TTGGTGACAT TTATTGGAAT ATGCAGTTAG ACTCACTCAG AGGGACTGGT ATAGATTTGA ACATACTTC CTGTGAACAA AGGAAATCGC TCGAAAGCCA TCAAAAATGTA AAAAATTATT TGTTGTCAAC TCTTCCCGAA TCAACAGGTA CTCTCCTGCC TCCATTAGAA | 3300 3360 3420 3480 3540 |
| 25 | TGCAACAAAT CAAGCCTTGA TAAATCCAAG ATAATCTCTA TTTCGCCCAA TCCGGCGAAA GCTGTTGTAA CAATAATCTA CTATACCGAT AACCCTTCCT GTTCTGTAAT AAAAATATAT GGAATAAATG GAGCCTCGGC TGATATAACC GGGTTGCCCA AACATCTATC CGAAGGTTAT TACAGCATAC AGTTCAATAC ATCCAACTTT GATCCGGAT TCTACCTGGT AACGCTAAAT GTTGATCAGA AAATTATAGA TACGGAAAAA TTACGAATCA AA | 3600 3660 3720 3780 3822 |
| 30 | (2) INFORMATION FOR SEQ ID NO:203 | |
| 35 | (i) SEQUENCE CHARACTERISTICS: (A) LEMGTH: 2775 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | |
| 40 | (ii) MOLECULE TYPE: DMA (genomic) | |
| | (iv) ANTI-SEMSE: NO | |
| 4 5 | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | |
| 50 | <pre>(ix) FEATURE: (A) UAIE/KEY: misc feature (B) LOCATION 12775</pre> | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:203 | |
| 55 | ATGGCTATCA TGATGAAAAG TATTGTTTTT AGAGCATTTC TAACGATTTT GCTCTGTGG GCAGCGATCA CGAATCCGAC TGCTCAAGAG ATCTCAGGCA TGAATGCATC CTGTCTGGCT GCTCCGGCTC AACCGGATAC TATCTTATAT GAAGTTTTG AGAATGGACC TGTTCCCAAT GGCTGGCTTG AGATGATGC TGATGCTGAT GGGGCACTT GGGGAAGCCC ATCAGGCTCT TTCTCTGTAC CTTACGGACA CAATGGCCTT TGCACCTACT CCCATATACG TTCCGGTATC | 60 120 180 240 300 |
| 60 | TCAACASCGG GCAACTATCT GATTACACCC AATATAGAAG GAGCCAAACG GGTCAAGTAC TGGGTATGCA ATCAGTATAG TACCAATCCG GAACATTACG CAGTAATGGT ATCGACAACG GGGACTGCCA TTGAAGACTT TGTTTTGTTG TTTGATGATT CCATAACAGG GAAACCGACT CCTCTTGTAT GGCGTAGACG AATCGTGGAC TTACCGGAAG GGACCAAATA TATTGCATGG | 360 420 480 540 |
| 65 | CGACATTACA AAGTCACCGA CTCACACACA GAATTCTTGA AATTGGATGA TGTCACTGTG TATAGGTCGA TCGAAGGGCC CGAACCTGCT ACCGACTTCA CAGTAATCAA TATTGGTCAG AATGTGGGAC GATTGACTTG GAACTATCCG GAGGATTATC AACCGGAAGG AAAGGGGAAT GAAGAGTTGC AGCTTAGCGG CTACAACATC TATGCGAACG GTACACTACT GGCACAAATA AAAGATGTCT CCATACTGGA GTATGTGGAC AGCACTTACT CTTTGCGAGA CAATCCCTTG | 600 660 720 780 840 |
| 70 | CAAGTGGAGT ACTGCGTTAC AGCCGTTTAC GATGAAAGCA TAGAATCTTC GACCGTATGT GGCACGCTGC ATTACGCCAC GGATGCCATC CTTTATGAAA ATTTTGAGAA TGGACCTGTT CCCAATGGTT GGCTTGTGAT AGACGCTGAT GGAGATGGAT TTAGCTGGGG ACACTATTTG AATGCATACG ACGCTTTTCC CGGCCATAAT GGAGGCCATT GCTCCTTGTC GGCTTCTTAT GTTCCGGGTA TAGGCCCGGT GACTCCCGAC AACTATCTGA TTACCCCCAA GGTTGAAGGA | 900 960 1020 1080 1140 |
| 75 | GCCAAACGTG TCAAGTACTG GGTAAGCACG CAGGATGCCA ATTGGGCAGC GGAACATTAC GCGGTGATGG CTTCGACAAC GGGGACTGCT GTCGGAGATT TCGTCATATT GTTCGAAGAA ACCATGACAG CGAAGCCGAC CGGCGCATGG TATGAAAGAA CCATCAACTT ACCTGAAGGG | 1200 1260 1320 |

WO 99/29870 PCT/AU98/01023

```
ACTAAATACA TCGCATGGCG GCATTACAAC TGTACCGATA TATATTTCTT GAAGTTGGAC
                                                                                    1380
       GATATCACTG TATTCGGGAC TCCTGCATCA GAGCCCGAAC CTGTTACCGA TTTCGTTGTC
       TCGCTTATTG AAAACAACAA GGGACGATTA AAGTGGAATT ATCCTAACGG CTACGAACCC
                                                                                    1500
       GATAAGACTG ATGATAAAGA CCCATTGCAG CTTGCCGGCT ACAATATCTA TGCAAACGGC
                                                                                    1560
 5
       TCGCTCCTTG TTCACATACA AGACCCGACT GTTTTGGAGT ATATCGATGA GACTTATTCT
                                                                                    1620
       TCACGAGACG ATCAGGTGGA AGTGGAATAT TGTGTCACTG CCGTTTATAA CGACAATATC
                                                                                    1680
       GAGTCCCAAT CGGTTTGCGA TAAGCTGATT TATGATTCTC AATCGGACAT TATCTTATAT GAAGGCTTTG AGGCCGGAAG TATTCCTGAA GGCTGGTTGT TGATTGATGC TGATGGCGAC
                                                                                    1740
                                                                                    1800
       AATGTTAATT GGGACTATTA TCCTTGGACT ATGTATGGAC ATGACAGTGA GAAGTGTATTA
GCATCCCCTT CGTACTTACC GATGATTGGC GTTTTAACTC CGGATAACTA TTTGGTTACA
                                                                                    1860
10
                                                                                    1920
       CCCAGACTCG AAGGAGCCAA GCTTGTCAAG TATTGGGTAA GTGCGCAAGA TGCTGTTTAT
                                                                                    1980
       TCGGCTGAGC ATTATGCTGT GATGGTTTCT ACTACGGGAA CTGCTGTTGA AGATTTTGTC
                                                                                    2040
       CTCTTGTTCG AAGAGACAAT GACCGCTAAG GCTAACGGTG CATGGTATGA GCGAACTATT
                                                                                    2100
       ACATTGCCTG CAGGAACAAA ATATATTGCC TGGCGGCATT ATGATTGCAC CGATATGTTT
                                                                                    2160
15
       TTCTTGCTCT TGGATGACAT TACGGTTTAT CGTTCTACTG AGACTGTTCC CGAGCCTGTT
                                                                                    2220
       ACTGATTTCG TTGTCTCGCT TATTGAGAAT AACAAGGGTC GCCTGAAATG GAATTATCCT
                                                                                    2280
       AACGGCTACG AACCCGATAA GACTGATGAT AAAAAACCAT TGCAGCTTAC CGGCTACAAC
                                                                                    2340
       ATCTATGCAA ATGGCTCGCT CCTTGTTCAC ATACAAGACC CGACTGTTTT GGAGTATATC
                                                                                    2400
       GATGAGACTT ATTCTTCACG AGACGGTCAG GTGGAAATGG AATATTGTGT CACTGCCGTT
                                                                                    2460
20
       TATAACGACA ATATCGAGTC CCAATCGGTT TGCGATAAGC TGAACTATAC TATCACATCC
TTGGATAATA TTCAATCTGA TACAAGCTTG AAAATATATC CTAATCCGGC ATCGTATGTG
                                                                                    2520
                                                                                    2580
       GTAAGGATAG AGGGATTGAG TCGGAGCAAG TCGACAATCG AGTTGTATAA TGCGCTGGGA
                                                                                    2640
       ATTTGCATAT TAAGGGAAGA GACTCATTCA GAGAAAACGG AAATCGATGT TTCACGTCTC
                                                                                    2700
       NATGACGGAG TCTACTTGAT TAAAGTAGTC GGTGGAAATA AAACAACAAC CGAAAAGGTA
                                                                                    2760
25
       GAGATAAAGA GGCCG
       (2) INFORMATION FOR SEQ ID NO: 204
30
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2766 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
35
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
40
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...2766
           (xi) SEQUENCE DESCRIPTION: SEO ID NO: 204
50
      ATGATGAAAA GTATTGTTTT TAGAGCATTT CTAACGATTT TGCTCTCGTG GGCAGCGATC
      ACGAATCCGA CTGCTCAAGA GATCTCAGGC ATGAATGCAT CCTGTCTGGC TGCTCCGGCT
                                                                                     120
      CAACCGGATA CTATCTTATA TGAAAGTTTT GAGAATGGAC CTGTTCCCAA TGGCTGGCTT
                                                                                     180
      GAGATAGATG CTGATGCTGA TGGTGCCACT TGGGGAAGCC CATCAGGCTC TTTCTCTGTA
                                                                                     240
55
      CCTTACGGAC ACAATGGCCT TTGCACCTAC TCCCATATAC GTTCCGGTAT CTCAACAGCG
                                                                                     300
      GGCAACTATC TGATTACACC CAATATAGAA GGAGCCAAAC GGGTCAAGTA CTGGGTATGC
                                                                                     360
      AATCAGTATA GTACCAATCC GGAACATTAC GCAGTAATGG TATCGACAAC GGGGACTGCC
ATTGAAGACT TTGTTTTGTT GTTTGATGAT TCCATAACAG GGAAACCGAC TCCTCTTGTA
                                                                                     420
                                                                                     480
      TGGCGTAGAC GAATCGTGGA CTTACCGGAA GGGACCAAAT ATATTGCATG GCGACATTAC
                                                                                     540
60
      AAAGTCACCG ACTCACACA AGAATTCTTG AAATTGGATG ATGTCACTGT GTATAGGTCG
                                                                                     600
      ATCGAAGGGC CCGAACCTGC TACCGACTTC ACAGTAATCA ATATTGGTCA GAATGTGGGA
                                                                                     660
      CGATTGACTT GGAACTATCC GGAGGATTAT CAACCGGAAG GAAAGGGGAA TGAAGAGTTG
                                                                                     720
      CAGCTTAGCG GCTACAACAT CTATGCGAAC GGTACACTAC TGGCACAAAT AAAAGATGTC
                                                                                    780
      TCCATACTGG AGTATGTGGA CAGCACTTAC TCTTTGCGAG ACAATCCCTT GCAAGTGGAG
                                                                                     840
65
      TACTGCGTTA CAGCCGTTTA CGATGAAAGC ATAGAATCTT CGACCGTATG TGGCACGCTG
                                                                                     900
      CATTACGCCA CGGATGCCAT CCTTTATGAA AATTTTGAGA ATGGACCTGT TCCCAATGGT
                                                                                     960
      TGGCTTGTGA TAGACGCTGA TGGAGATGGA TTTAGCTGGG GACACTATTT GAATGCATAC
                                                                                   1020
      GACGCTTTTC CCGGCCATAA TGGAGGCCAT TGCTCCTTGT CGGCTTCTTA TGTTCCGGGT
                                                                                   1080
      ATAGGCCCGG TGACTCCCGA CAACTATCTG ATTACCCCCA AGGTTGAAGG AGCCAAACGT
                                                                                   1140
70
      GTCAAGTACT GGGTAAGCAC GCAGGATGCC AATTGGGCAG CGGAACATTA CGCGGTGATG
                                                                                   1200
      GCTTCGACAA CGGGGACTGC TGTCGGAGAT TTCGTCATAT TGTTCGAAGA AACCATGACA
                                                                                   1260
      GCGAAGCCGA CCGGCGCATG GTATGAAAGA ACCATCAACT TACCTGAAGG GACTAAATAC
ATCGCATGGC GGCATTACAA CTGTACCGAT ATATATTTCT TGAAGTTGGA CGATATCACT
                                                                                   1320
                                                                                   1380
      GTATTCGGGA CTCCTGCATC AGAGCCCGAA CCTGTTACCG ATTTCGTTGT CTCGCTTATT
                                                                                   1440
75
      GAAAACAACA AGGGACGATT AAAGTGGAAT TATCCTAACG GCTACGAACC CGATAAGACT
                                                                                   1500
```

PCT/AU98/01023

1560

140/490

```
GTTCACATAC AAGACCCGAC TGTTTTGGAG TATATCGATG AGACTTATTC TTCACGAGAC
                                                                               1620
       GATCAGGTGG AAGTGGAATA TTGTGTCACT GCCGTTTATA ACGACAATAT CGAGTCCCAA
                                                                               1680
       TCGGTTTGCG ATAAGCTGAT TTATGATTCT CAATCGGACA TTATCTTATA TGAAGGCTTT
                                                                               1740
       GAGGCCGGAA GTATTCCTGA AGGCTGGTTG TTGATTGATG CTGATGGCGA CAATGTTAAT
                                                                               1800
       TGGGACTATT ATCCTTGGAC TATGTATGGA CATGACAGTG AGAAGTGTAT TGCATCCCCT
                                                                               1860
       TCGTACTTAC CGATGATTGG CGTTTTAACT CCGGATAACT ATTTGGTTAC ACCCAGACTC
                                                                               1920
       GAAGGAGCCA AGCTTGTCAA GTATTGGGTA AGTGCGCAAG ATGCTGTTTA TTCGGCTGAG
                                                                               1980
       CATTATGCTG TGATGGTTTC TACTACGGGA ACTGCTGTTG AAGATTTTGT CCTCTTGTTC
                                                                               2040
10
       GAAGAGACAA TGACCGCTAA GGCTAACGGT GCATGGTATG AGCGAACTAT TACATTGCCT
                                                                               2100
       GCAGGAACAA AATATTTGC CTGGCGGCAT TATGATTGCA CCGATATGTT TTTCTTGCTC
                                                                               2160
       TTGGATGACA TTACGGTTTA TCGTTCTACT GAGACTGTTC CCGAGCCTGT TACTGATTTC
                                                                               2220
       GTTGTCTCGC TTATTGAGAA TAACAAGGGT CGCCTGAAAT GGAATTATCC TAACGGCTAC
                                                                               2280
       GAACCCGATA AGACTGATGA TAAAAAACCA TTGCAGCTTA CCGGCTACAA CATCTATGCA
                                                                               2340
15
       AATGGCTCGC TCCTTGTTCA CATACAAGAC CCGACTGTTT TGGAGTATAT CGATGAGACT
                                                                               2400
       TATTCTTCAC GAGACGGTCA GGTGGAAATG GAATATTGTG TCACTGCCGT TTATAACGAC
                                                                               2460
       AATATCGAGT CCCAATCGGT TTGCGATAAG CTGAACTATA CTATCACATC CTTGGATAAT
                                                                               2520
      ATTCAATCTG ATACAAGCTT GAAAATATAT CCTAATCCGG CATCGTATGT GGTAAGGATA
                                                                               2580
       GAGGGATTGA GTCGGAGCAA GTCGACAATC GAGTTGTATA ATGCGCTGGG AATTTGCATA
                                                                               2640
20
      TTAAGGGAAG AGACTCATTC AGAGAAAACG GAAATCGATG TTTCACGTCT CAATGACGGA
                                                                               2700
      GTCTACTTGA TTAAAGTAGT CGGTGGAAAT AAAACAACAA CCGAAAAGGT AGAGATAAAG
                                                                               2760
      AGGCCG
                                                                               2766
25
       (2) INFORMATION FOR SEQ ID NO: 205
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2763 base pairs
                 (B) TYPE: nucleic acid
30
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
35
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
40
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...2763
45
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205
      ATGAAAAGTA TTGTTTTTAG AGCATTTCTA ACGATTTTGC TCTCGTGGGC AGCGATCACG
      AATCCGACTG CTCAAGAGAT CTCAGGCATG AATGCATCCT GTCTGGCTGC TCCGGCTCAA
                                                                                120
50
      CCGGATACTA TCTTATATGA AAGTTTTGAG AATGGACCTG TTCCCAATGG CTGGCTTGAG
                                                                                180
      ATAGATGCTG ATGCTGATGG TGCCACTTGG GGAAGCCCAT CAGGCTCTTT CTCTGTACCT
                                                                                240
      TACGGACACA ATGGCCTTTG CACCTACTCC CATATACGTT CCGGTATCTC AACAGCGGGC
      AACTATCTGA TTACACCCAA TATAGAAGGA GCCAAACGGG TCAAGTACTG GGTATGCAAT
      CAGTATAGTA CCAATCCGGA ACATTACGCA GTAATGGTAT CGACAACGGG GACTGCCATT
                                                                                420
55
      GAAGACTTTG TTTTGTTT TGATGATTCC ATAACAGGGA AACCGACTCC TCTTGTATGG
CGTAGACGAA TCGTGGACTT ACCGGAAGGG ACCAAATATA TTGCATGGCG ACATTACAAA
                                                                                480
                                                                                540
      GTCACCGACT CACACAGA ATTCTTGAAA TTGGATGATG TCACTGTGTA TAGGTCGATC
      GAAGGGCCCG AACCTGCTAC CGACTTCACA GTAATCAATA TTGGTCAGAA TGTGGGACGA
                                                                                660
      TTGACTTGGA ACTATCCGGA GGATTATCAA CCGGAAGGAA AGGGGAATGA AGAGTTGCAG
                                                                                720
60
      CTTAGCGGCT ACAACATCTA TGCGAACGGT ACACTACTGG CACAAATAAA AGATGTCTCC
      ATACTGGAGT ATGTGGACAG CACTTACTCT TTGCGAGACA ATCCCTTGCA AGTGGAGTAC
                                                                                840
      TGCGTTACAG CCGTTTACGA TGAAAGCATA GAATCTTCGA CCGTATGTGG CACGCTGCAT
                                                                                900
      TACGCCACGG ATGCCATCCT TTATGAAAAT TTTGAGAATG GACCTGTTCC CAATGGTTGG
                                                                                960
      CTTGTGATAG ACGCTGATGG AGATGGATTT AGCTGGGGAC ACTATTTGAA TGCATACGAC
                                                                               1020
65
      GCTTTTCCCG GCCATAATGG AGGCCATTGC TCCTTGTCGG CTTCTTATGT TCCGGGTATA
                                                                               1080
      GGCCCGGTGA CTCCCGACAA CTATCTGATT ACCCCCAAGG TTGAAGGAGC CAAACGTGTC
                                                                               1140
      AAGTACTGGG TAAGCACGCA GGATGCCAAT TGGGCAGCGG AACATTACGC GGTGATGGCT
TCGACAACGG GGACTGCTGT CGGAGATTTC GTCATATTGT TCGAAGAAAC CATGACAGCG
                                                                               1200
                                                                               1260
      AAGCCGACCG GCGCATGGTA TGAAAGAACC ATCAACTTAC CTGAAGGGAC TAAATACATC
                                                                               1320
70
      GCATGGCGGC ATTACAACTG TACCGATATA TATTTCTTGA AGTTGGACGA TATCACTGTA
                                                                               1380
      TTCGGGACTC CTGCATCAGA GCCCGAACCT GTTACCGATT TCGTTGTCTC GCTTATTGAA
                                                                               1440
      AACAACAAGG GACGATTAAA GTGGAATTAT CCTAACGGCT ACGAACCCGA TAAGACTGAT
                                                                               1500
      GATAAAGACC CATTGCAGCT TGCCGGCTAC AATATCTATG CAAACGGCTC GCTCCTTGTT
                                                                               1560
      CACATACAAG ACCCGACTGT TTTGGAGTAT ATCGATGAGA CTTATTCTTC ACGAGACGAT
                                                                               1620
75
      CAGGTGGAAG TGGAATATTG TGTCACTGCC GTTTATAACG ACAATATCGA GTCCCAATCG
                                                                               1680
```

GATGATAAAG ACCCATTGCA GCTTGCCGGC TACAATATCT ATGCAAACGG CTCGCTCCTT

PCT/AU98/01023

```
GTTTGCGATA AGCTGATTTA TGATTCTCAA TCGGACATTA TCTTATATGA AGGCTTTGAG
                                                                               1740
                                                                               1800
      GCCGGAAGTA TTCCTGAAGG CTGGTTGTTG ATTGATGCTG ATGGCGACAA TGTTAATTGG
      GACTATTATC CTTGGACTAT GTATGGACAT GACAGTGAGA AGTGTATTGC ATCCCCTTCG
                                                                               1860
      TACTTACCGA TGATTGGCGT TTTAACTCCG GATAACTATT TGGTTACACC CAGACTCGAA
                                                                               1920
      GGAGCCAAGC TTGTCAAGTA TTGGGTAAGT GCGCAAGATG CTGTTTATTC GGCTGAGCAT
 5
                                                                               1980
      TATGCTGTGA TGGTTTCTAC TACGGGAACT GCTGTTGAAG ATTTTGTCCT CTTGTTCGAA GAGACAATGA CCGCTAAGGC TAACGGTGCA TGGTATGAGC GAACTATTAC ATTGCCTGCA
                                                                               2040
                                                                               2100
                                                                               2160
      GGAACAAAAT ATATTGCCTG GCGGCATTAT GATTGCACCG ATATGTTTTT CTTGCTCTTG
                                                                               2220
      GATGACATTA CGGTTTATCG TTCTACTGAG ACTGTTCCCG AGCCTGTTAC TGATTTCGTT
10
                                                                               2280
      GTCTCGCTTA TTGAGAATAA CAAGGGTCGC CTGAAATGGA ATTATCCTAA CGGCTACGAA
      CCCGATAAGA CTGATGATAA AAAACCATTG CAGCTTACCG GCTACAACAT CTATGCAAAT
                                                                               2340
      GGCTCGCTCC TTGTTCACAT ACAAGACCCG ACTGTTTTGG AGTATATCGA TGAGACTTAT
                                                                               2400
      TCTTCACGAG ACGGTCAGGT GGAAATGGAA TATTGTGTCA CTGCCGTTTA TAACGACAAT
                                                                               2460
      ATCGAGTCCC AATCGGTTTG CGATAAGCTG AACTATACTA TCACATCCTT GGATAATATT
                                                                               2520
15
      CAATCTGATA CAAGCTTGAA AATATATCCT AATCCGGCAT CGTATGTGGT AAGGATAGAG
                                                                               2580
      GGATTGAGTC GGAGCAAGTC GACAATCGAG TTGTATAATG CGCTGGGAAT TTGCATATTA
                                                                               2640
                                                                               2700
      AGGGAAGAGA CTCATTCAGA GAAAACGGAA ATCGATGTTT CACGTCTCAA TGACGGAGTC
                                                                               2760
      TACTTGATTA AAGTAGTCGG TGGAAATAAA ACAACAACCG AAAAGGTAGA GATAAAGAGG
      CCG
                                                                               2763
20
      (2) INFORMATION FOR SEQ ID NO: 206
            (i) SEQUENCE CHARACTERISTICS:
25
                 (A) LENGTH: 1779 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
30
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
35
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
40
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1779
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:206
45
                                                                                 60
      ATGAACAGCA TCATGAAATA TCAATTATAT ACGGCCGTCA TAATGGCTCT CTCTGTATCA
      TCCGTTTGCG GTCAAACCC ACGAAATACA GAAACCAAAC GCCCCGACAC GCTGCGCAGG
                                                                                120
      GAGCTTACTA TCGTTAATGA CCAGACTGTG GAGATGGAGC ATGCGGATCC GCTTCCGGCT
                                                                                180
      GCATACAAGG CCATCGAACC TCGATTAAAA CCTTTCCGTC CGGAATATAA CAAGCGTACA
                                                                                240
      TTCGGATTTG TCCCTGAAGT TTCCTCTTCA GGCAGGAACA ATCTTCCGAA TATCCTGCCG
                                                                                300
50
      ACGGAAGGTC ATATGAAGCA CCGGGGGTAC CTGAATATCG GTATCGGCCA TACGCTAAAC
                                                                                360
      CAGCGAATGG ATGCCGGCTA TCGTCTGATA GATGCAGAGC AGGAGAGACT GAATCTTTTC
                                                                                420
                                                                                480
      CTCTCCTATC GTGGGATGAA ATCGGCTTTC AATACCGGTG ACTTCGACGG CGACAGAAAG
      GATAGACGAA TGATGGCAGG AGTGGACTAC GAGCAGCGCA GGCCTTCCTT TGTGCTTGCT
                                                                                540
      ACCGGCTTGT ATTATTCGAA CCATTATTC AATAACTACG GACGGGAGC TACCACCAAT
                                                                                600
55
      GTGGGCAGCA TCCCTCAGCT ATCGACACCT GTTACTCCTC AGATGGACAA CGGGACCCAC
                                                                                720
      AACGTCCGTG TATACTTGGG TGCAAAAAAT GATGTGATCG ATGCCAGGAT CGACTATCGT
      TTCTTCCGTT CTATTCCCTA TCTGGGTACC GATCCGATGA AGGCTCTCAC AGAACATACG
                                                                                780
      CCTGAACTGA ACGTGACGAT GAGTAATGAG TTGTCCGATG ATATTAAGCT CGGTGTCGAA GTTCGTACGG GAGGATTGTT TTTTGCCAAA AACAGCGAAA TGATTCAAAC GGGCGTTCTG
                                                                                840
                                                                                900
60
                                                                                960
      TCCGAAACCG ACCGCAACCT GTATTATGTG GAGGGCGCGC CCACAATCGG ATTTGTCGGA
                                                                               1020
      GACTCGGACA ATATGCAATG GAACATACAG GCCGGAGTAG GGATTTCTTC CCATTTCGGA
      GCCAAAGGGA GGTTGTTTTT CTGGCCTAAA CTGGATGCTT CGCTTAGTAT CTTCCCTTCA
                                                                               1080
      TGGCGTGTGT ATGCGAAAGC CTTCGGCGGT GTGATTCGAA ATGGTCTCGC CGATGTTATG
                                                                               1140
      CAAGAGGAGA TGCCCTACCT GATGCCCAAT ACGATTGTAC TCCCTTCGCG CAATGCTTTG
                                                                               1200
65
      ACCGCCCAAT TAGGGGTGAA GGGGAATATA GCCGATGTGG TACGTATGGA GGTTTATGGC
                                                                               1260
      GACTICICCA AGCIGACAGG IGIGCCITIC TATACICCGA CICTACCCII ATATAAICCA
                                                                               1320
      TCCGACTTGT ATCAGTATAA TGTGAGTTTC TTGCCGATAT ATGCCGACGG CAGCCGCTGG
                                                                               1380
      CGCGCAGGTG GTAAGCTGGA ATACTCTTAT CGCGATATGC TCCGCTTTCT GGTAGACGCA
                                                                               1440
      TCCTATGGCA AGTGGAATTT GGATGGAGGA CTTGTCGCCT CCATGCAGCC CGATCTTATA
70
      TTGAAGGCAG AAGTAGGTGT TCATCCCATT GCCCCATTGG ATGTCAGACT CCGGTATACA
                                                                               1560
                                                                               1620
      CAGCTGAACG GACGGTATCG GTATTCTTTC GGCTCGGCTG GCTCGGAAGC CTTGGGTATC
      GGTAATGTAC ATCTTCTTAG TGCGGATGTT TCATACAAGC TGAAAAAAAA CTTGAGCCTT
                                                                               1680
      TATCTCAAAA TCGATAATAT GCTGGCGGAA ACGACAGAGC TTATCGGTTA TTATCCTATG
                                                                               1740
      CAGCCGTTCC ATTGTTTCGC CGGTTTTAGC TGGACTTTC
                                                                               1779
75
```

```
(2) INFORMATION FOR SEQ ID NO: 207
            (i) SEQUENCE CHARACTERISTICS:
 5
                  (A) LENGTH: 1767 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
10
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
15
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
20
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1767
           (xi) SEQUENCE DESCRIPTION: SEO ID NO:207
25
       ATGAAATATC AATTATATAC GGCCGTCATA ATGGCTCTCT CTGTATCATC CGTTTGCGGT
       CAAACCCCAC GAAATACAGA AACCAAACGC CCCGACACGC TGCGCAGGGA GCTTACTATC
GTTAATGACC AGACTGTGGA GATGGAGCAT GCGGATCCGC TTCCGGCTGC ATACAAGGCC
                                                                                      120
                                                                                      180
       ATCGAACCTC GATTAAAACC TTTCCGTCCG GAATATAACA AGCGTACATT CGGATTTGTC
                                                                                      240
       CCTGAAGTTT CCTCTTCAGG CAGGAACAAT CTTCCGAATA TCCTGCCGAC GGAAGGTCAT
                                                                                      300
30
       ATGAAGCACC GGGGGTACCT GAATATCGGT ATCGGCCATA CGCTAAACCA GCGAATGGAT
                                                                                      360
       GCCGGCTATC GTCTGATAGA TGCAGAGCAG GAGAGACTGA ATCTTTTCCT CTCCTATCGT
                                                                                      420
       GGGATGAAAT CGGCTTTCAA TACCGGTGAC TTCGACGGCG ACAGAAAGGA TAGACGAATG
                                                                                      480
       ATGGCAGGAG TGGACTACGA GCAGCGCAGG CCTTCCTTTG TGCTTGCTAC CGGCTTGTAT
                                                                                      540
       TATTCGAACC ATTATTCAA TAACTACGGA CGGGGAGCTA CCACCAATGT GGGCAGCATC
                                                                                      600
35
       CCTCAGCTAT CGACACCTGT TACTCCTCAG ATGGACAACG GGACCCACAA CGTCCGTGTA
                                                                                      660
       TACTTGGGTG CAAAAATGA TGTGATCGAT GCCAGGATCG ACTATCGTTT CTTCCGTTCT
                                                                                      720
       780
       GTGACGATGA GTAATGAGTT GTCCGATGAT ATTAAGCTCG GTGTCGAAGT TCGTACGGGA
       GGATTGTTT TTGCCAAAAA CAGCGAAATG ATTCAAACGG GCGTTCTGTC CGAAACCGAC
                                                                                      900
40
       CGCAACCTGT ATTATGTGGA GGGCGCCC ACAATCGGAT TTGTCGGAGA CTCGGACAAT
                                                                                      960
       ATGCAATGGA ACATACAGGC CGGAGTAGGG ATTTCTTCCC ATTTCGGAGC CAAAGGGAGG
                                                                                    1020
       TTGTTTTTCT GGCCTAAACT GGATGCTTCG CTTAGTATCT TCCCTTCATG GCGTGTGTAT
                                                                                    1080
       GCGAAAGCCT TCGGCGGTGT GATTCGAAAT GGTCTCGCCG ATGTTATGCA AGAGGAGATG
CCCTACCTGA TGCCCAATAC GATTGTACTC CCTTCGCGCA ATGCTTTGAC CGCCCAATTA
                                                                                    1140
                                                                                    1200
45
       GGGGTGAAGG GGAATATAGC CGATGTGGTA CGTATGGAGG TTTATGGCGA CTTCTCCAAG
CTGACAGGTG TGCCTTCTA TACTCCGACT CTACCCTTAT ATAATCCATC CGACTTGTAT
                                                                                    1260
                                                                                    1320
       CAGTATAATG TGAGTTTCTT GCCGATATAT GCCGACGGCA GCCGCTGGCG CGCAGGTGGT
AAGCTGGAAT ACTCTTATCG CGATATGCTC CGCTTTCTGG TAGACGCATC CTATGGCAAG
                                                                                    1380
                                                                                    1440
       TGGAATTTIGG ATGGAGGACT TGTCGCCTCC ATGCAGCCCG ATCTTATATT GAAGGCAGAA GTAGGTGTTC ATCCCATTGC CCCATTGGAT GTCAGACTCC GGTATACACA GCTGAACGGA
                                                                                    1500
50
                                                                                    1560
       CGGTATCGGT ATTCTTTCGG CTCGGCTGGC TCGGAAGCCT TGGGTATCGG TAATGTACAT
                                                                                    1620
       CTTCTTAGTG CGGATGTTTC ATACAAGCTG AAAAAGAACT TGAGCCTTTA TCTCAAAATC
                                                                                    1680
       GATAATATGC TGGCGGAAAC GACAGAGCTT ATCGGTTATT ATCCTATGCA GCCGTTCCAT
                                                                                    1740
       TGTTTCGCCG GTTTTAGCTG GACTTTC
                                                                                    1767
55
       (2) INFORMATION FOR SEQ ID NO:208
            (i) SEQUENCE CHARACTERISTICS:
60
                  (A) LENGTH: 1038 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
65
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SENSE: NO
70
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
75
                  (A) NAME/KEY: misc_feature
```

(B) LOCATION 1...1038

| (xi) | SEQUENCE | DESCRIPTION: | SEQ | ΙD | 110:208 |
|------|----------|--------------|-----|----|---------|
|------|----------|--------------|-----|----|---------|

| _ | | | | | | | |
|----|------------|--------------|--------------|------------|------------|------------|------|
| 5 | ATGATGGAAA | AATGTATCTT | TGCTCACTAT | CCACATAACC | TTGTGTTCAT | GATTCGCAAG | 60 |
| | CATTTCGGTA | TCATTTTGGG | ATTTCTTTCT | CTTGTGTTTT | CGGCAGGTGC | TCAACAAGAG | 120 |
| | AAGCAGGTGT | TTCATTTTCT | GAACCTTCCG | GCTACTGCAC | AGGCTTTGGC | TGCCGGAGGC | 180 |
| | AAAGCTATCA | CCATCGTAGA | CGACAATCCC | GGACTGGCTT | TTGAGAATCC | GGCTCTGCTC | 240 |
| | GGATATGAAT | CCGGTGGCCG | CGCCTTTCTT | TCCTATTTAT | ATTATATGAG | TGGTTCGCAT | 300 |
| 10 | ATGGGCAATG | CCTGTTATGC | CTCGTCCGTC | GGAGAGCGTG | GCATGTGGGG | TGTTGGCATG | 360 |
| | CGTTTCCTGA | ACTACGGGTC | TATGCAAGGA | TACGATCAGA | ATGCGATTGC | CACCGGCTCT | 420 |
| | TTTAGTGCTT | CGGATATAGC | TGTACAAGGA | TTTTACAGCC | ATGAACTGAG | CAACCACTTC | 480 |
| | CGCGGTGGAG | TCAGCCTAAA | AGCATTGTAT | TCTTCTATCG | AGACGTATAG | TTCCTTTGGC | 540 |
| | CTTGGTGTGG | ATGTCGGTAT | CAGTTATTAC | GACGATGACA | AAGGATATTC | CGCTTCCGCT | 600 |
| 15 | CTGTTCAAGA | ACGTAGGGGC | GCAACTGAAA | GGCTATAATG | AAGAACGGGA | ACCGCTCGAT | 660 |
| | TGGGATTTCC | AGCTCGGCTT | TTCCCGCAGT | TTTATCAATG | CTCCGTTTCG | CTTGCACATC | 720 |
| | ACGTTGTTCA | ATCTGAATCC | GCACTATTTC | AAGCGTCTTG | TACCACGCGA | TCTGTCCAAG | 780 |
| | ATGCAAAAGT | TCCTCCGACA | CTTCTCGATA | GGAGCAGAAT | TTACTCCTTC | CGAGAGGTTT | 840 |
| | TGGGTCGGGC | TGGGATATAC | GCCACAGATT | GCACAGGATT | TCGAGGTGGA | AGGCGGCAAC | 900 |
| 20 | AAATGGGGAG | GTCTTTCGGC | CGGCGTCGGT | TTCACTTCAG | GTGTAGTACG | TGTAGGCGTA | 960 |
| | TCTGCTGCCA | CCTATCATCC | TGCAGCTCTT | TCGTTCATGT | GTTCGGTAGG | TATCCGTTTG | 1020 |
| | GACGATAAGA | GCATCTTC | | | | | 1038 |
| | | | | | | | |
| ~- | | | | | | | |
| 25 | (2) INFORM | ATION FOR SE | EQ ID NO:209 | 9 | | | |
| | | | | | | | |

- - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1035 base pairs
- (B) TYPE: nucleic acid 30
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
 - (ii) MOLECULE TYPE: DNA (genomic)
- 35 (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- 40 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 - (ix) FEATURE:
 - (A) NAME/KEY: misc feature
 - (B) LOCATION 1...1035
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209

| ATGGAAAAAT | GTATCTTTGC | TCACTATCCA | CATAACCTTG | TGTTCATGAT | TCGCAAGCAT | 60 |
|------------|--|---|---|---|--|--|
| TTCGGTATCA | TTTTGGGATT | TCTTTCTCTT | GTGTTTTCGG | CAGGTGCTCA | ACAAGAGAAG | 120 |
| CAGGTGTTTC | ATTTTCTGAA | CCTTCCGGCT | ACTGCACAGG | CTTTGGCTGC | CGGAGGCAAA | 180 |
| GCTATCACCA | TCGTAGACGA | CAATCCCGGA | CTGGCTTTTG | AGAATCCGGC | TCTGCTCGGA | 240 |
| TATGAATCCG | GTGGCCGCGC | CTTTCTTTCC | TATTTATATT | ATATGAGTGG | TTCGCATATG | 300 |
| GGCAATGCCT | GTTATGCCTC | GTCCGTCGGA | GAGCGTGGCA | TGTGGGGTGT | TGGCATGCGT | 360 |
| TTCCTGAACT | ACGGGTCTAT | GCAAGGATAC | GATCAGAATG | CGATTGCCAC | CGGCTCTTTT | 420 |
| AGTGCTTCGG | ATATAGCTGT | ACAAGGATTT | TACAGCCATG | AACTGAGCAA | CCACTTCCGC | 480 |
| GGTGGAGTCA | GCCTAAAAGC | ATTGTATTCT | TCTATCGAGA | CGTATAGTTC | CTTTGGCCTT | 540 |
| GGTGTGGATG | TCGGTATCAG | TTATTACGAC | GATGACAAAG | GATATTCCGC | TTCCGCTCTG | 600 |
| TTCAAGAACG | TAGGGGCGCA | ACTGAAAGGC | TATAATGAAG | AACGGGAACC | GCTCGATTGG | 660 |
| GATTTCCAGC | TCGGCTTTTC | CCGCAGTTTT | ATCAATGCTC | CGTTTCGCTT | GCACATCACG | 720 |
| TTGTTCAATC | TGAATCCGCA | CTATTTCAAG | CGTCTTGTAC | CACGCGATCT | GTCCAAGATG | 780 |
| CAAAAGTTCC | TCCGACACTT | CTCGATAGGA | GCAGAATTTA | CTCCTTCCGA | GAGGTTTTGG | 840 |
| GTCGGGCTGG | GATATACGCC | ACAGATTGCA | CAGGATTTCG | AGGTGGAAGG | CGGCAACAAA | 900 |
| TGGGGAGGTC | TTTCGGCCGG | CGTCGGTTTC | ACTTCAGGTG | TAGTACGTGT | AGGCGTATCT | 960 |
| GCTGCCACCT | ATCATCCTGC | AGCTCTTTCG | TTCATGTGTT | CGGTAGGTAT | CCGTTTGGAC | 1020 |
| GATAAGAGCA | TCTTC | | | | | 1035 |
| | TTCGGTATCA CAGGTGTTTC GCTATCACCA TATGAATCCG GGCAATGCCT TTCCTGAACT AGTGCTTCGG GGTGGATCA GGTGTGGATG TTCCAAGAACG GATTTCCAGC TTGTTCAATC CAAAAGTTCC GTCGGGCTCG GTGGGAGTC GCTGCCACCT | TTCGGTATCA TTTTGGGATT CAGGTGTTTC ATTTTCTGAA GCTATCACCA TCGTAGACGA TATGAATCCG GTGGCCGCG GGCAATGCCT GTTATGCCTC TTCCTGAACT ACGGGTCTAT AGTGCTTCGG ATATAGCTGT GGTGGAGTCA GCCTAAAAGC TTCAAGAACG TAGGGGCGCA GATTTCCAGC TCGGCTTTC TTGTTCAATC TGAATCCGCA CAAAAGTTCC TCCGACACTT GTGGGGTGG GCTGCCACCT ATCAGCCGG GCTGCCACCT ATCAGCCGG | TTCGGTATCA TTTTGGGATT TCTTTCTTT CAGGTGTTTC ATTTCTGAA CCTTCCGGCT GCTATCACCA TCGTAGACGA CAATCCCGGA TATGAATCCG GTGGCCGCC CTTTCTTTCC GGCAATGCCT GTTATGCCTC GTCCGTCGGA TCCTGAACT ACGGGTCTAT GCAAGGATTC AGTGCTTCGG ATATAGCTGT ACAAGGATTT GGTGTGGATC TCGGTATCAG TTATTACGAC TTCAAGAACG TAGGGGCGCA ACTGAAAGGC GATTTCCAGC TCGGCTTTC CCGCAGTTTT TTGTTCAATC TGAATCCGCA CTATTTCAAG CAAAAGTTCC TCCGACACTT CTCGATAGGA TCGGGCTGG GATATACGC ACAGATTGCA TGGGGGGGTC TTTCGGCCGG CTCTGTTTC GCTGCCACCT ATCATCCTC AGCTCTTTCG | TTCGGTATCA TTTTGGGATT TCTTTCTTT GTGTTTTCGG CAGGTGTTTC ATTTCTGAA CCTTCCGGCT ACTGCACAGG GCTATCACCA TCGTAGACGA CAATCCCGGA CTGGCTTTTG TATGAATCCG GTGGCCGCC CTTTCTTTCC TATTTATATT GGCAATGCCT GTTATGCCTC GTCCGTCGGA GAGCGTGGCA TTCCTGAACT ACGGGTCTAT GCAAGGATTC TACAGCATG GGTGAGGTCA GCCTAAAAGC ATTGTATTCT TCTATCGAGA GGTGTGGATG TCGGTATCAG TTATTACAC GATGACAAAG TTCAAGAACG TAGGGGCGCA ACTGAAAGGC TATAATGAAG GATTTCCAGC TCGGCTTTC CCGCAGTTTT ATCAATGCTC TTGTTCAATC TCGAACCTC CTGATAGGA GCAGAATTTA CTCGGGCTGG GATATACGCC ACAGATTGCA CAGGATTTCA TCGGGCTGG GATATACGCC ACAGATTCA CAGGATTTCA GTCGGGCTGG GATATACGCC ACAGATTCCA ACAGATTCCA GCTGCCACCT ATCATCCTGC ACCTCTTTC ACTTCAGGTG GCTGCCACCT ATCATCCTGC AGCTCTTTC ACTTCAGGTG | TTCGGTATCA TTTTGGGATT TCTTTCTCTT GTGTTTTCGG CAGGTGCTCA CAGGTGTTTC CAGGTGTTTC CAGGTGTTTC CAGGTGTTTC CAGGTGTTTC CATTTCTGAA CCTTCCGGCT CAGGTGTTTC CAGGTGAACACA CCTTCCGGCT CAGGTGTTTTG CAGGATGCCC GTTATGCCTC GTCGTCGGA CATCCCGGA CATCCCGGA CATGCACACA CATGCACACA CATTTCCTGAA CACGGTCTAT CACAGGATC CACAGGATC CAGGTGTACAC CACAGGATC CACAGGATC CACAGGATC CACAGGATC CACAGGATC CACAGGATC CACAGGATC CACAGGATC CACAGATC CACAGATTC CACAGATC CACAGATC CACAGATTC CACAGATC CACAGATTC CACAGATT CACAGAT C | CAGGTGTTTC ATTTCTGAA CCTTCCGGCT ACTGCACAGG CTTTGGCTGC CGGAGGCAAA GCTATCACCA TCGTAGACGA CAATCCCGGA CTGCTTTTG AGAATCCGGC TCTGCTCGGA TATGAATCCG GTGGCGCGC CTTTTCTTTCT TATTTATATT ATTAGATGG TTCGCATATGGGCAATCCGGT TTCCTGAACT ACGGGTCTAT GCAAGGATAC GATCAGAATG CGATTGCCAC CGGCTCTTTT AGTGCTCG ATATAGCTGT ACAAGGATAC GATCAGAATG CGATTGCCAC CGGCTCTTTT ACAGCCATG ACTGAGACT ACGGGTCAAAAGC ATTGTATTCT TCTATCGAGA CATTGCCAC CTTTTGGCCTT GGTGTGGATG TCGGTATCAG ATTGTATTCT TCTATCGAGA CATTGCCAC CTTTTGGCCTT TTCAAGAACG TAGGGGCGCA ACTGAAAAGC TATAATGAACG TAGGGGCCA ACTGAAAAGG TATAATGAACG TCGGATTTC CCGCAGTTTT ATCAATGCTC CGTTTCGCTT GCACATCACG TTGTTCAATC TCCGACACTT ATCAATGCTC CACGCGATCT GTCCAAGAATG CACAGAATGCA ACAGGATTCG ACAGAATTTC CACGCGACTCT GTCCAAGAAG GATGATACGCC ACAGATTCAG ACAGAATTTC CTCTTCCGA GAGGTTTTGGCCACCT TCCGCACACACA ACAGATTCACG ACAGATTCA ACAGATTCA ACAGATTCA ACAGATTCA ACAGATTCA ACAGATTCAG CGCCAACAAA TGGGGGGGTC TTTCGGCCG CGTCGGTTTC ACTTCAGGTG TAGTACGTC AGGCTACTAG AGGGTACT AGCGTAACACACACACACACACACACACACACACACACAC |

- (2) INFORMATION FOR SEQ ID NO:210
- 70 (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 990 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRAHDEDNESS: double (D) TOPOLOGY: circular
- 75

45

```
(ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
 5
           (iv) AUTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
10
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...990
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:210
15
      ATGATTCGCA AGCATTTCGG TATCATTTTG GGATTTCTTT CTCTTGTGTT TTCGGCAGGT
      GCTCAACAAG AGAAGCAGGT GTTTCATTTT CTGAACCTTC CGGCTACTGC ACAGGCTTTG
                                                                                 120
      GCTGCCGGAG GCAAAGCTAT CACCATCGTA GACGACAATC CCGGACTGGC TTTTGAGAAT
                                                                                 180
      CCGGCTCTGC TCGGATATGA ATCCGGTGGC CGCGCCTTTC TTTCCTATTT ATATTATATG
                                                                                 240
20
      AGTGGTTCGC ATATGGGCAA TGCCTGTTAT GCCTCGTCCG TCGGAGAGCG TGGCATGTGG
                                                                                 300
      GGTGTTGGCA TGCGTTTCCT GAACTACGGG TCTATGCAAG GATACGATCA GAATGCGATT
GCCACCGGCT CTTTTAGTGC TTCGGATATA GCTGTACAAG GATTTTACAG CCATGAACTG
                                                                                 360
                                                                                 420
      AGCAACCACT TCCGCGGTGG AGTCAGCCTA AAAGCATTGT ATTCTTCTAT CGAGACGTAT
                                                                                 480
      AGTTCCTTTG GCCTTGGTGT GGATGTCGGT ATCAGTTATT ACGACGATGA CAAAGGATAT
                                                                                 540
25
      TCCGCTTCCG CTCTGTTCAA GAACGTAGGG GCGCAACTGA AAGGCTATAA TGAAGAACGG
                                                                                 600
      GAACCGCTCG ATTGGGATTT CCAGCTCGGC TTTTCCCGCA GTTTTATCAA TGCTCCGTTT
                                                                                 660
      CGCTTGCACA TCACGTTGTT CAATCTGAAT CCGCACTATT TCAAGCGTCT TGTACCACGC
                                                                                 720
      GATCTGTCCA AGATGCAAAA GTTCCTCCGA CACTTCTCGA TAGGAGCAGA ATTTACTCCT
                                                                                 780
      TCCGAGAGGT TTTGGGTCGG GCTGGGATAT ACGCCACAGA TTGCACAGGA TTTCGAGGTG
                                                                                 840
30
      GAAGGCGCA ACAAATGGGG AGGTCTTTCG GCCGGCGTCG GTTTCACTTC AGGTGTAGTA
                                                                                 900
      CGTGTAGGCG TATCTGCTGC CACCTATCAT CCTGCAGCTC TTTCGTTCAT GTGTTCGGTA
                                                                                 960
      GGTATCCGTT TGGACGATAA GAGCATCTTC
                                                                                 990
35
      (2) INFORMATION FOR SEQ ID NO:211
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 972 base pairs
                 (B) TYPE: nucleic acid
40
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
45
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
50
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...972
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:211
      ATGTGCCTCG AACCCATAAT TGCTCCGATT TCATCCGAGT TGCTCGAGCA GGAGCTGACT
      GCCGATCGTT TTCTGCGGAT GACAAACAAA GCCGGCAATG AGATCTATGT TTTTACGGCC
                                                                                 120
60
      GAAGAAGCTC CGCATTGCAT GAAAGAAGTA GGCCGACTGC GAGAAGAAGC CTTTCGGCAT
                                                                                 180
      TATGGCGGAG GTACTGGCAA GGCGATCGAT ATAGACGAGT TCGACACCAT GCCCGGGAGC
                                                                                 240
      TACAAACAGC TGATCGTATG GGATCCGCAA AACAAGGCTA TACTCGGAGG CTACCGCTTT
                                                                                 300
      ATCTATGGGC GGGACGTTGC TTTCGATACC GATGGCAAGC CTTTGCTGGC AACGGCAGAG ATGTTTCGCT TCAGTGATGC TTTTTTGCAC GATTATCTCC CCTACACAGT CGAATTGGAA
                                                                                 360
                                                                                 420
65
      CGTTCGTTCG TGTCGCTCCA GTACCAATCG ACACGGATGG GCACAAAGGC CATTTTTGTG
                                                                                 480
      CTGGACAATC TTTGGGACGG TATCGGAGCA CTCACTGTAG TCAATCCAGA GGCACTCTAT
                                                                                 540
      TTCTATGGCA AGGTGACCAT GTACAAAGAC TATGATCGGC GAGCTCGCAA TCTGATCCTG
                                                                                 600
      TATTTTCTTC GCAAGCACTT CTCCGATCCG GAAGGCTTGG TCAAGCCTAT TCATCCCCTA
                                                                                 660
      CCGATAGAGA TCAGTGCGGA GGACGAAGCC TTGTTCTCCT CATCCGACTT TGACACCAAT
                                                                                 720
70
      TACAAGACTC TCAATATAGA AGTGCGCAAG CTGGGTATCA ATATCCCTCC TCTCGTGAGT
                                                                                 780
      GCATATATAG CTTTGTCTCC GGAGATGCGT GTTTTCGGCA CTGCAGTGAA TGAGTCTTTC
                                                                                 840
      GGAGAGGTGG AGGAAACCGG CATATTCATT GCTGTGGGTA AGATCCTGGA AGAGAAAAAA
                                                                                 900
      CAACGGCACA TAGAGAGCTT CATCCTCAGC CGGAACGAAA AAAAAGGTCT CGACAGTAGC
                                                                                 960
      AATGGCCGAT CA
                                                                                 972
75
```

```
(2) INFORMATION FOR SEQ ID NO:212
            (3) SEQUENCE CHARACTERISTICS:
 5
                  (A) LENGTH: 1641 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
10
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) AUTT-SEUSE: NO
15
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
20
                  (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1641
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:212
25
      ATGAAAACCA TTGTAAGATA CAGCCGCCTT CCGGTCGCTC TCTTCTTTTG CCTTTTTGGGA
      GCTGTGCACT TGTCTGTCGA GGCGCAGATG CTCAATACTC CTTTCGAGCT GTCGGATCAG
                                                                                  120
      ATCGTCCTGT CTCCCACCGA AAGGCAATAC AGGGAGATTT GTGTGCAAAC GAAAGAAAAA
                                                                                  180
      AGGGGGGCCG ATCTTTTCCC GTTGAGCGAT AAGCTGCGCG ATTCGGCCTA TGTTCGTTTC
                                                                                  240
      GGCTCGGCCT ATGGCGATAT TGCGGGCGAC TATCTTCCGT ACAACGGCAA TAACTACTCC
                                                                                  300
30
      TCGCTCTCGC TCGAATCGGG TGGTCGCATC AGTGTCCGTA ACTATGGCAC ATTGCAGGGC
                                                                                  360
      AGTGCTTCCT ACTCACGTGG CATGCACAAA CGCATCGGCT GGAATGCTCT GCGCAACGCC
      GAAGCCTACT ATCCCTATTT GGTGTCCGAT TCGACCGGCG GAGACTATCA TTTCGAAGAC
                                                                                  480
      TATCGGCTTG CCGGCTACTA TTCTTTTCGC GCCGGCCGCT TGCCCCTCGG TATAGGCTTC
                                                                                  540
      TCATACAGGG GCGAAGTTGC TTATCGGCTG ACCGATCCGC GTACGACCAA TACGACCGGT
                                                                                  600
35
      GCATTGGAGC TTTCTTGTGC TACCTCTTTG ACGCTGCCTC GAGAGAACAG GCTATCGCTT
                                                                                  660
      TCGGCTGCGT ATCTCTATCA TAGACAACAC CTCACACAGT ACAACTGGCG TCCCGGGCAG
                                                                                  720
      CAGGACAAAT TCTTCGTCAG CTACGGTTTC GGTCAGGTGG ATGTCAGCAA CAGCCCTATC
                                                                                  780
      TGGTTCGGTA TCTCCAGAAT GAACTACGTC AACGGATGGA AGCTTAGCTC CCGTCTGGAT ACCCGTAGGG GCGATGCCAT CGGTCTCGAC TACAGCGGCT ACTTCCTCGA TACCGAGAG
                                                                                  840
                                                                                  900
      AGGTCGTCCA TCAATCTCTT TGCTTTGCTT TACAATCGCC TGCGACTCTA TGGTAGCTGG
CATCTGTCGG ACTTCGATTT TTCATTTCA GCCGACTATG CTCTGCGCCA AGGGATAGAG
40
                                                                                  960
                                                                                 1020
      CGGATATACG AAGACTACAA GCCGGATGAT AATTATCATA TCTACGACCT CCGTATCTTG
                                                                                 1080
      GCCATTCGCC GCTGGTATAT GCTCAATGAG TTTTCTGCCC AAGCCCAAGC CTCCTACCGT
                                                                                 1140
      ATTCGCACGG ATAGAGGTTG TGCCCTGAGA GTGAGTGCCG GTAGTGATTT CTACGGCTAT
                                                                                 1200
45
      GATGAGACGT ATCGCAAGCA TGGACATCAT ACCATGAGCG GAATGCTACG TCCTTTTGCC
                                                                                 1260
      GGTATAGCCT ATGACCATGC CGGATCCAAA TTGGATTTTG GACTTTCGCT TTCGGCTGCT
      TATCGAATGG TGCTGACGCA TTCGTATAAG ATTCGTACCA TCCAGAAAGA GCAGCTCGAC
                                                                                 1380
      TATCAGCTGG CCTATTTGCC CTATGCCTAT CGTAATAGAG AAGGCGTGGA GGTGCGTTCC
                                                                                 1440
      TCTCTGTACG TCTCGATTCC GATGCAGAAT ACCCACCGCC TGATGACAGA GCTGCGGTTG
                                                                                 1500
50
      TATGGCGACC TGATGAAAAG AAAGGACGGT ATAGCCTATG GCAAAACGCC CGGTGTCATC
                                                                                 1560
      TCACATATCC TGTCCGATCC GCAAGCCGAA CGAACGTCCG GCCATACCAT CGGGGCTATC
                                                                                 1620
      TGCAATATCT CCTACCTCTT C
                                                                                 1641
55
      (2) IMFORMATION FOR SEQ ID NO:213
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2250 base pairs
                  (B) TYPE: nucleic acid
60
                 (C) STRAHDEDHESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
65
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
70
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (E) LOCATION 1...2250
75
```

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213
      ATGAAAAAC TTCACATGAT TGCCGCCTTA GCCGTCCTGC CTTTCTGCCT GACGGCACAA
                                                                              60
      GCACCCGTCT CCAACAGCGA GATAGATAGT CTTAGCAATG TGCAGCTCCA GACCGTACAG
                                                                             120
 5
      GTCGTAGCTA CTCGCGCCAC GGCGAAAACC CCTGTCGCTT ACACCAACGT TCGCAAGGCC
                                                                             180
      GAACTTTCCA AGTCCAATTA TGGTCGTGAC ATCCCCTATC TGCTGATGCT GACTCCCTCC
                                                                             240
      GTGGTAGCCA CCAGCGATGC CGGTACGGGT ATCGGATATT CCGGCTTTCG CGTGCGTGGC
                                                                             300
      ACCGATGCCA ATCGCATCAA CATAACTACC AATGGAGTAC CCCTCAACGA CTCCGAATCT
                                                                             360
      CASTCCGTCT TTTGGGTGAA TATGCCCGAC TTCGCCTCTT CCATCGAAGA CCTTCAGGTG
                                                                             420
10
      CAGCGAGGTG TGGGTACTTC CACCAATGGT GCCGGAGCTT TTGGGGCAAG TGTCAATATG
                                                                             480
      CGTACGGATA ATTTGGGACT GGCTCCTTAT GGCCGTGTCG ATTTGAGCGG AGGTTCGTTC
                                                                             540
      GGCACATTCC GCCGATCGGT CAAACTCGGT AGCGGACGCA TCGGTCGCCA TTGGGCAGTG
                                                                             600
      GATGCCCGCC TGTCCAAAAT CGGTTCGGAC GGCTACGTGG ATAGAGGAAG CGTGGATCTG
                                                                             660
      AAATCCTATT TCGCACAGGT GGGCTATTTC GGTAGCAACA CGGCTCTCAG GTTCATCACT
                                                                             720
15
      TTCGGAGGAA AAGAAGTTAC GGGTATCGCA TGGAACGGTC TTTCCAAGGA GGATGAAGCC
                                                                             780
      AAATATGGCC GCCGATACAA CAGTGCCGGT CTTATGTACG TGGACGCGCA AGGAGTACCG
                                                                             840
      CACTACTACC ACAATACCGA CAATTACGAG CAGCGTCACT ACCATGCCAT CATGACGCAC
                                                                             900
      AGCTTCTCTC CTTCCGTTAT CCTCAACCTC ACGGCACACT ACACGGCCGG ATATGGCTAT
                                                                             960
      ACGGACGAAT ATCGTACCGG ACGTAAACTA AAGGAATATG CACTGCAGCC CTATGTGGAA
                                                                            1020
20
      AACAGTGTGA CCGTGAAGAA AACGGATCTC ATCCGTCAGA AGTATCTGGA CAATGACTTC
                                                                            1080
      GGAGGACTCA TCGGTTCGCT TAACTGGCAC ACCGGTGCAT GGGATTTGCA GTTCGGGGCC
                                                                            1140
      TCGGGCAATA TCTATAAAGG AGACCACTTC GGCCGTATCA CTTACATCAA AAAGTACAAT
                                                                            1200
      CAGCCCTTAG CTCCCGACTT CGAATATTAT CGGAACAGGG CAGACAAAAG AGAAGGTGCA
                                                                            1260
      GCCTTTGCCA AAGCCAACTG GCAGATCACT CCGGAACTGA ACATGTATGC CGACCTCCAG
                                                                            1320
25
      TATOGTACCA TOGGCTACAC GATAAACGGC ATCACGGACG AATATGATGA GGTACAGGGA
      AG!ATGCAGC ACATCGATTT GGACAAGACC TTCCGCTTCC TCAATCCGAA GGCCGGTCTT
                                                                            1440
      ACCTATAGTT TCGACGATGC TCATACTGCC TATGCTTCTG TTGCGGTAGC ACACCGCGAG
                                                                            1500
      CCTAACAGAA CCAATTACAC CGAAGCCGGA ATAGGACAGT ATCCTACGCC TGAGCGACTG
                                                                            1560
      ATCGACTATG AGCTGGGCTA CCGCTATGCT TCGCCCCTCT TGTCGGCCGG AGTAGGTCTC
                                                                            1620
30
      TATTATATGC AATACAAGGA CCAACTCGTG CTGGATGGCC GTTTGAGCGA TGTGGGACAG
                                                                            1680
      ATGCTCACAA GCAACGTCCC CGACAGCTAC CGTATGGGAC TGGAGCTGAC TCTCGGTTGG
                                                                            1740
      CAGATCCTTC CTCGTTTGCT GCGTTGGGAT GCTTCTTTCA CTATGAGTCG CAACAAAATC
                                                                            1800
      GACCGCTACG TACAATATAC ATCCGTATAT GATGCGGACT ACAACTGGCT CGAACTCAAG
                                                                            1860
      GAGGAGACCC TCGAAAGCAC GGATATAGCC TACTCGCCCA ATGTCATTGC CGGCAGCATG
                                                                            1920
35
      CTTACCCTCT CTCATGCCGG TTTCGAAATG GCTTGGACGA GCCGCTTCGT CAGCAAGCAA
                                                                            1980
      TATCTGGACA ATACACAGCG CAGCGATCGC ATGCTTTCCT CCTATTGGGT GAACGACCTC
                                                                            2040
      CGCCTCGGCT ATGTGCTGCC GGTTCACTTC GTTAAGAGAG TGGCACTGGG CGTACAGCTC
                                                                            2100
      AATAATCTCT TCAACCTCAT GTATGCGTCC AATGCCTACA TCTACGATGC CGGTTACGTA
                                                                            2160
      CAGGCATCCG GAGAACTAAG TGCATATGCC GATCTGCGTT ATTATCCTCA GGCCGGATTT
                                                                            2220
40
      AATGCACTGG GTAGTCTGAC AATCGATTTC
                                                                            2250
      (2) INFORMATION FOR SEQ ID NO:214
45
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1482 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
50
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
55
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISH: PORTPHYROMONAS GINGIVALIS
60
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...1482
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:214
65
      ATGAAAAGAA GGTTTCTATC GCTGTTACTG CTGTACATAC TCTCTTCCAT CAGCCTTTCT
      GCTCAGCGGT TTCCGATGGT GCAGGGAATC GAGTTGGATA CCGATTCGCT TTTCTCTCTG
                                                                             120
      CCCAAGCGTC CTTGGCGCGC CATCGGTAAA ACGATAGGCG TCAATCTGGC CGTATGGGGC
                                                                             180
      TTCGATCATT TCATCATGAA CGAGGACTTT GCAGACATCA GTTGGCAGAC TATCAAGAGC
                                                                             240
70
      AATTTCCAAA CAGGCTTTGG CTGGGACAAT GACAAGTTTG TCACCAACCT CTTCGCACAT
                                                                             300
      CCTTATCACG GATCGCTCTA TTTCAATGCA GCGAGGTCGA ACGGTTTGAG CTTCAGGCAC
                                                                             360
      TCTGCTCCGT TTGCCTTCTT TGGCAGTCTC ATGTGGGAGC TGCTTATGGA AAACGAGCCA
                                                                             420
      CCGAGTATCA ACGACCTCTG TGCCACCACC ATAGGCGGTA TAGCTTTGGG GGAGATGGGG
                                                                             480
      CACAGGCTGT CGGACCTGCT CATCGACAAT CGTACCACAG GGTGGGAACG TATGGGGCGC
                                                                             540
75
      GASGTGSCTA TCGCTCTGAT CAATCCGATG CGCTTTCTCA ACCGTCTGAC AGCAGGASAG
                                                                             600
```

```
GTGACTTCTG TCGGGAGTCG CAGCGGACAG ATATTTCAGT CTGTCCCCAT AAACATAGTC GTCGATGCCG GCTTTCGCTT TTTGGCAGAC AAGCGGCATG CCCGAACCGG TGCCACGGCT
                                                                                  660
                                                                                  720
       CTGACCCTGA ATCTGAGATT CGACTACGGC GATCCATTCC GAAGCGAGAC TTTCTCTCCA
                                                                                  780
       TACGATTCT TCCAATTCAA AGCCGGATTG AGTTTCTCCG AATCGCAACC TCTGCTGAGC
                                                                                  840
 5
       CAGATCAATC TGATCGGAAT CCTAAGCGGA TGCCAACTGC TCGCACACGA ACGAACGGTT
                                                                                  900
       TTGGTGGGAG GTCTCTTCA GCACTTCGAC TACTACAATT CGGAAAAACG AATAAGCAAA
                                                                                  960
       AATTCGGAGG AGGTACTCGT CACCCCATAC CGTATCTCGC AAGTGGCAGC TCTGGGAGGC
                                                                                1020
       GGTCTTATCT TCCAGCACCA CGGAAAATTT CGACGACGTC CTCTGGAGCT ATATGCCGAG
                                                                                1080
       ACCTACCTGA ATGTCGTCCC GATGGGAGCC AGTCTGTCGG ATCACTACAA CGTGGACAAT
                                                                                1140
10
       CGGGACTATA ACCTCGGCAG CGGATTGAGC GGCAAGCTAT ACCTTGGTGC TACGTACAAT
                                                                                1200
       GATCTGTGGA GCTGGCTCTT GGGAGTCGAA AGCTATCGGC TCTACACATG GATCGGGTAT
                                                                                1260
       GAAGAGCCGC ACCAGAAAAA TACCGATGTC AGCTCTTTTA TGGTGCAGGG GGACGAAAGC
                                                                                1320
      AAGGCGCCC TACTGGTGAC GAGTTCCGAG TTCGCATTTC ATCCTGGCCC CTGGCATGTA
                                                                                1380
      GCCATCGTCG CTCGCCGTTT CATCCGCAAA ACAGCCTATC AATTCTACCC TAACGTATCA
                                                                                1440
15
      TTCGATACCG GCGACATACA GCTGCGTGTC GGATTTCACT TC
                                                                                1482
       (2) INFORMATION FOR SEQ ID NO:215
20
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 882 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
25
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
30
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROHOMAS GINGIVALIS
35
           (ix) FEATURE:
                  (A) MAME/KEY: misc_feature
                  (B) LOCATION 1...882
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:215
40
      ATGAAACGAC TGATTGTTTT TCTGGCAATG GGTGGCTTGC TGTTCACCCT TGCGAACGCA
                                                                                  60
      CAAGAAGCAA ACACTGCATC TGACACTCCC AAAAAGGACT GGACTATAAA AGGTGTGACC GGACTAAATG CCTCTCAGAC TTCTCTGACC AACTGGGCTG CCGGTGGAGA AAACACGGTG
                                                                                 120
                                                                                 180
      GCAGGTAACC TCTATTTGAA CATAGATGCC AACTACCTGA AAGATAAATG GAGTTGGGAC
                                                                                 240
45
      AACGGTTTGC GTACAGACTT CGGTCTGACC TACACAACAG CCAACAAGTG GAACAAAAGT
                                                                                 300
      GTAGACAGA TCGAACTCTT CACGAAGGCC GGCTATGAGA TCGGCAAACA TTGGTACGGA
      AGTGCGCTTT TCACTTTCCT CTCACAGTAT GCCAAAGGAT ATGAGAAGCC CTCGGATCAC
                                                                                 420
      TTGACAGGAG TCAAGCATAT CTCTAATTTC TTCGCTCCTG CATATCTCAC TCTCGGTATT
                                                                                 480
      GGTGCGGACT ATAAGCCCAA TGAGAAGTTC TCTCTCTACC TCTCTCCTAC AACGGGCAAG
                                                                                 540
50
      CTGACTGTAG TAGCAGACGA CTACCTCTCA AGTTTGGGAG CCTTCGGGGT GAAAGTTGGT
                                                                                 600
      GAAAAGACAA TGTTCGAACT TGGTGCTTTG GTAGTGGGTT CGGCCAATAT AAATCTGATG
                                                                                 660
      GAGAATGTCA ATTTGATAAC CAAGGCTTCA TTCTTCTCGG CTTATACGCA CGACTTTGGC
                                                                                 720
      AACATTGACA TCAATTGGGA GGCTATGCTG GCCATGAAGA TCAACAAGTT CCTCACGGCT
                                                                                 780
      ACGATAGCCA CCAATCTTAT CTACGACGAT GATGTGAAGA TCAACGATGG CCCGAAAATC
                                                                                 840
55
      CAGTTCAAAG AAGTTGTGGG CGTGGGTGTT GCGTACACTT TC
                                                                                 882
      (2) INFORMATION FOR SEQ ID NO:216
60
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 612 base pairs
                  (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
65
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
70
          (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
75
           (ix) FEATURE:
```

```
(A) NAME/KEY: misc feature
                  (B) LOCATION 1...6\overline{1}2
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:216
  5
       ATGAAGAAA TGATTTTGGC AGCTACTATG CTGCTCGCAA CAATCGGTTT TGCAAATGCT
       CAGAGTCGTC CTGCTCTTAG ACTGGATGCT AACTTTGTCG GTAGTAACTT AATGCAAAAA
                                                                               120
       GTCGCAAACA CGAGCGTGAA CAATAAGATG ATCGTAGGCT TACGTGTTGG TGCTGCTGCT
                                                                               180
       GAGTTCGCTC TTAGCAATGA TGGATTCTAT CTCGCCCCCG GATTGGCCTA TACGATGAGA
                                                                               240
10
       GGTGCTAAGA TGGAATCACT AAGTGAAACG ACAACTCGCT TGCATTATCT GCAAATACCG
                                                                               300
       GTGAATGCCG GTATGAGATT TAGCTTTGCT GACAACATGG CTATTTCATT GGAAGCAGGT
                                                                               360
       CCCTATTTCG CATATGGTGT CGCCGGAACG ATTAAGACTA AAGTTGCAGG CGTTACGGCT
                                                                               420
       TCTGTAGATG CCTTTGGTGA TAACGGATAT AACCGTTTCG ACTTGGGCTT GGGCTTGTCT
                                                                               480
       GCTGCCTTGA GCTACGACCG TTATTACGTA CAAATTGGAT ATGAGCATGG ATTGCTTAAT
                                                                               540
15
       ATGTTGAAGG ATGCTCCGGA TAAGACTTCT TTGCGTAATC ATGACTTCTT TGTGGGTCTC
                                                                               600
       GGTGTTCGCT TC
                                                                               612
       (2) INFORMATION FOR SEQ ID NO:217
20
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 729 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
25
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
30
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMOMAS GINGIVALIS
35
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...729
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:217
      ATGAAAAGGA TTTTTACTGT AGCCCTTGTG CTACTTGCTT CGGTCACTAT GGCCATCGGA
                                                                               60
       CAAAGCCGCC CGGCACTTCG CGTAGATGCC AACTTCGTAG GCAGCAATCA GAGCATGAAA
                                                                              120
       AGAGACGGAT ATGTGTGGGA CACCAAAATG AATGTCGGCC TGCGGGTCGG TGCCGCTGCC
                                                                              180
45
       GAATTCATGA TCGGATCAAG AGGATTCTAC TTGGCTCCGG GTCTGAACTA TACGATGAAG
                                                                              240
       GGCTCCAAAA CCGAATGGGA TATACCCGAA ATGGTTCCTG GTACCTATAT TACGATGGTT
                                                                              300
       TCCACTCGCT TGCACTATCT GCAACTGCCG ATCAATGCCG GCATGCGGTT CGACCTGATG
      AATGACATGG CGGTTTCGAT CGAAGCGGGT CCTTTCCTTG CATACGGTAT ATATGGTACA
                                                                              420
      TATCGGCAGA AGTTGGAAGG ATGGAAGCCG AACAACTACA GCACAGAGTT TTTTGGCCCA
                                                                              480
50
      ACSCTTGGTG GCCCAACAAA TATCCGCTGG GACATCGGGG CAAACATAAT AGCCGCATTC
                                                                              540
      CACTATAAGC GTTATTATAT ACAGATAGGC TATGAACATG GATTTGTGGA TATTGTGTCA
                                                                              600
      GGTGGAGGTT CTGATATTCC CCGACTGAAC GACAATAGGC AATCCTCTTC GACGACCGCT
                                                                              660
      CTAAGAGAAA AGGGAAATAA CGAATACGCT TATAATCGTG ACTTCTTCGT GGGCATAGGT
                                                                              720
      TACCGCTTT
                                                                              729
55
      (2) INFORMATION FOR SEQ ID NO:218
            (i) SEQUENCE CHARACTERISTICS:
60
                 (A) LENGTH: 621 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
65
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
70
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
75
                (A) NAME/KEY: misc_feature
```

(B) LOCATION 1...621

| (xi) | SEQUENCE | DESCRIPTION: | SEQ | ID | 110:218 |
|------|----------|--------------|-----|----|---------|
|------|----------|--------------|-----|----|---------|

| 5 | ATGAAAAGAA | TGCTGCTGCT | TCTCGTTGTA | TTATTATATG | GAATTGCAGG | CCGATTGGCT | 60 |
|----|------------|------------|------------|------------|------------|------------|-----|
| | GCACAAGACG | TTATCAGACC | ATGGTCATTG | CAGGTCGGAG | CGGGATACTC | CGATACGGAG | 120 |
| | AACATCCCGG | GAGGATTCAC | CTATGGTTTC | TATTTGGGAA | AGCGTATGGG | GAGCTTTCTG | 180 |
| | GAAGTGGGGC | TGTCCATGTA | CAACTCCACA | CGTCAAACAG | CCAACAATGC | AGACTCCTTT | 240 |
| | GCATCGAACG | AAGGAGACGG | ATCTTTTCAG | GTAAATATGT | CTTCTCCGAA | TGAGAAGTGG | 300 |
| 10 | TCATTCTTCG | ATGCAGGCAG | TGCCAACTGC | TATATGATCG | TCGTCGGAGT | CAATCCTCTC | 360 |
| | CATCTGTTTT | GGCAGAATAG | CCGGCACAAT | TTGTTTCTGG | CAGTACAAGC | CGGCCTGTCC | 420 |
| | AATAAGCACA | ATATTCATTT | CATCTATGGA | GACAAGGGAG | CCAAAGTCAG | TATCTACACC | 480 |
| | AATTCGAATA | CCTACATCGG | TTACGGAGCA | CGTGTAGCCT | ACGAATATCA | AATTCATAAA | 540 |
| | AACGTGGGGG | CGGGTGCCGC | TGTAATGTAC | GACCACGGCA | ATAAGATGCT | TACGGCCATG | 600 |
| 15 | GCCACGCTCT | CCACTCATTT | T | | | | 621 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO:219

| 20 | (i) SEQUENCE CHARACTERISTICS: |
|----|-------------------------------|
| | (A) LENGTH: 2853 base pairs |
| | (B) TYPE: nucleic acid |
| | (C) STRANDEDNESS: double |
| | (D) TOPOLOGY: circular |
| 25 | |

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 30 (iv) ANTI-SENSE: NO
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
- 35 (ix) FEATURE:
 - (A) NAME/KEY: misc feature
 - (B) LOCATION 1...2853
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:219

| 40 | | | | | | | |
|-----|------------|------------|------------|------------|------------|------------|------|
| | | | | TCATGGTTAT | TACCCGTAAT | CCTTGTCGGA | 60 |
| | TTGCTCTGTG | CTACTTTGGT | CGCTGCGGAA | CGTCCTATGG | CCGGAGCAGT | CGGATTGCAC | 120 |
| | CACCGTCGGC | ATGCTGCGCT | GTCTGATTCT | ACAGCGAAAG | ACACGGTGCC | TCTCGCAAAA | 180 |
| 4- | | | TCGAGATTCC | | ATTCCACCGG | ATCGATGCGG | 240 |
| 45 | | | TGAATTCGAA | TTGGAAGATA | TAGTGGAGTA | CGAAGCTGCC | 300 |
| | GATTCCATCG | TTTTGCTCGG | ACAGAATCGT | GCCTATCTTT | TCGGCAAGAG | CTATGTGAGC | 360 |
| | TATCAAAAGA | GTCGCTTGGA | GGCAAACTTC | ATGTATCTCA | ATACCGACAG | CAGTACGGTT | 420 |
| | TATACTCGCT | ATGTCCTCGA | TACGGCCGGT | TATCCGATGG | CCTTTCCTGT | TTTCAAGGAT | 480 |
| | GGAGAGCAGT | CGTTCGAAGC | CAAGAACTTT | ACCTACAACT | TCCGCACGGA | GAAGGGGATT | 540 |
| 50 | ATCAGCGGAG | TGATCACGCA | GCAGGGCGAA | GGCTATCTGA | CTGCCGGTAA | GACCAAGAAG | 600 |
| | ATGCCCGACA | ATATCATGTT | TATGCAAGGA | GGGCGTTATA | CGACCTGCGA | CAATCACGAT | 660 |
| | | | | GCAAAGGTGC | ATCCGGAGAA | AGACATCGTC | 720 |
| | | TCAATCTGGT | | ATGCCGCTGC | CGATAGGTCT | TCCTTTCGGC | 780 |
| | | TTTCCAACAA | | GGTATATTGA | TGCCCACGTA | CGGAGAGGAC | 840 |
| 55 | | GATTTTATTT | | GGATATTATT | | CGACTATATC | 900 |
| | | | GATCTTTTCC | AAAGGGTCAT | GGGGCATTTC | AGCCCAATCG | 960 |
| | | AGAGGTATAA | | | CCAATTATCT | GGTATCGAAG | 1020 |
| | | | | | CCAGTCTGAA | TATCCGATGG | 1080 |
| ~~ | | | GGCCAATCCT | | TGTCGGCCAA | | 1140 |
| 60 | | | | AATACCACCT | ATGATGTCAA | TGCCCGTACT | 1200 |
| | | GAAGTTCGGC | | | TTCCGGGTAC | | 1260 |
| | | | | ATGCGCGATA | CGACGGTGAG | CCTTACCTTG | 1320 |
| | | CGATTAATAT | | TATCCTTTCA | AGCGGAAGAC | CCGTGTAGGA | 1380 |
| G E | | GGTACGAGAA | | | GTCAGCTTCG | | 1440 |
| 65 | | AGAAAGATTT | | | GCGATTGGAA | | 1500 |
| | | TACCGATCAG | | | ATTATATCAA | | 1560 |
| | | | | | GGAAGTCGTG | | 1620 |
| | | TCCTGCCTTC | | | GCAGACTGTA | | 1680 |
| 70 | | | | | AGCCTTGGAA | | 1740 |
| 70 | | ATCTCATTAT | | | CCACTGTCAG | | 1800 |
| | | TCACGAAACG | | | TTCTTGAGCA | | 1860 |
| | | TGCATACGCT | | | AGCAGATATT | | 1920 |
| | | ATGCAGGATC | | | ACAACTTAGA | | 1980 |
| 75 | | | | | TGATAGATCA | | 2040 |
| 75 | TCTACATCCT | ATAATATGTT | TGCCGATTCG | ATCCGATGGA | GCAATATCTC | GGCTTCGCTG | 2100 |

```
GCACTTCGCC TCTCCAAGAG CTTTACCTTG CGCTTGTCCG GTCTGTTCGA TCCCTATTTG
                                                                                        2160
       ACGAAGTATT ATGAGGGAGA AGATGGGAAG ATCATTCCCT ATAAGAGCAA CGACCTGCGC
                                                                                        2220
       ATTTTTAACG GCAAGGGATT GGCACGCCTG ATCAGTACGG GTACTTCTTT CAGCTATACG
                                                                                        2280
       CTCAACAAAG AGTCGCTCAG CGGATTGATA GCTCTTTTCA GTGGCAAAAA GGAGCGGAGA
                                                                                        2340
 5
        GATGAAAAGA AAAACACAGG GGCTACTCCT CATGAAGGAG ACGATGCTGC CGATATACTT
                                                                                        2400
       GAGGGAGGAA GACCGCAAAA TGAAAGTGGG GGGTCGCTCC TCGAGCGCAA CCGTCAGGGC
                                                                                        2460
        GGAGCAGTGG ATCAGGATGG TTACTTCGCA TATTCGATCC CATGGAGCCT GTCCTTCGAC
                                                                                        2520
       TATAGTTGGA ATATTGCTAC CGACTACAAT AGGTACAATG TCAATAAGAT GGAGCACTAC
                                                                                        2580
       TACCGGGTAA CGCAGAATCT GAGCTTTCGC GGCAATATCC AGCCTACACC GAACTGGAGC
                                                                                        2640
10
       TTCGGATTCA ATGCGAACTA CAATTTCGAC TTGAAGAAAA TAACATCGCT TACCTGCAAC
                                                                                        2700
       GTCACTCGCG ACATGCACTG CTGGGCTATC TCGGCCAGTT TCATCCCTAT AGGAGCATAC AAGTCCTATA ATTTCGTCAT ATCGGTGAAG AGTTCACTCT TGCAGGATCT GAAGTATCAG
                                                                                        2760
                                                                                        2820
       CAGAGCAATC GTCCCATCAC GAATACTTGG TAT
                                                                                        2853
15
       (2) INFORMATION FOR SEQ ID NO: 220
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 3678 base pairs
20
                   (B) TYPE: nucleic acid
                   (C) STRAHDEDNESS: double
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
25
           (iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
30
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
35
                   (B) LOCATION 1...3678
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:220
       ATGATGAAAC GATATACAAT AATTCTTGCA GTTTTTCTTT TATTCTGCAC GGTATTTACC
                                                                                          60
40
       TTTCAAATAA AAGCTCGCCC TTATGAAAGA TTTGCAGATG TAGAGAAGCC TTGGATTCAG
                                                                                         120
       AAACATTCAA TGGATTCTAA ATTGGTGCCT GCAAATAAGG GTAACTTAAT TCAAGCTGAA
                                                                                         180
       ATTGTATACC AATCTGTTTC TGAACATAGT GACTTAGTTA TTTCACCTGT GAACGAAATA
       AGGCCTGCAA ATCGTTTCCC TTCGCATAGG AAGTCTTTTT TTGCAGAAAA TCTACGGGCA
       TCTCCCCCG TAGTTCCCGT TGCCGTCGAC AAGTATGCGG TACCGGTTGC CAATCCAATG
GATCCTGAAA ATCCCAATGC CTGGGATGTG ACGCTAAAAA TCACTACTAA AGCGGTAACA
45
                                                                                         420
       GTACCTGTCG ATGTGGTGAT GGTTATCGAC CAGTCTTCGT CAATGGGAGG GCAAAACATT
       GCCAGATTAA AGTCTGCCAT TGCATCGGGA CAGCGTTTTG TGAAAAAAAT GTTGCCTAAG
                                                                                         540
       GGGACGGCTA CAGAAGGGGT GCGTATCGCT CTTGTGAGTT ATGACCATGA GCCTCATCGC
                                                                                         600
       TTATCTGATT TTACCAAAGA CACTGCTTTT CTCTGTCAAA AAATCCGGGC TTTGACTCCT
                                                                                         660
50
       ATTTGGGGAA CACATACCCA GGGGGGGCTT AAAATGGCGA GAAACATTAT GGCCACTTCT
                                                                                         720
       ACTGCTGTGG ATAAGCATAT CATATTGATG TCTGACGGGT TAGCGACGGA GCAGTATCCT
                                                                                         780
       GTTAAAAATG TAACTACTGC AGACTTCATT GGCAAAACTG GAAATGCGAA TGATCCCATT
                                                                                         840
       GATTTGGTTA TACAAGGAGC AATTAATTTC CCTACAAATT ATGTTTCCAA CAATCCATCT ACACCTCTTA CCCCAAATTA TCCAACTCAT TCTTCTAAAG TTGGACGGAG AAATCTGCCG
                                                                                         900
                                                                                         960
55
       GAATCCAAAT TCGATTATAG TAATCTGAGT GCAAGGATTA CTTTTGATGG TGTTGCTGGC
                                                                                        1020
       GCATTGGTCT ATGAACCGAG GTTTCCTCAT CCCTATTATT ATTATTTCCC TTGTAACGCT
                                                                                        1080
       GCTATCAATG AGGCTCAGTT TGCGAAAAAC TCTGGTTATA CAATCCATAC TATTGGCTAT GACCTGGGAG ATTTTGCCTT GGCCAACAAT TCGTTGAAAC TAACCGCTAC AGACGAGAAT
                                                                                        1140
                                                                                        1200
       CACTTCTTTA CGGCGACACC GGCCAATTTA GCTGCAGCGT TTGATAATAT TGCCCAAACT
                                                                                        1260
60
       ATTAATATAG GTATACAGAG GGGGGAGGTG ACGGACTTTG TAGCTCCTGG TTTCATCGTT
                                                                                        1320
       AAAAATCTGA CGCAATCGGG AGATGTTACT CATTTGCTAA ATGTTTCAAA TGGAACGGTG
                                                                                        1380
       CACTATGATG TCTCTACTAA AAAACTGACA TGGACTACTG GTACTATCCT GAGCTCATCA
                                                                                        1440
       GAAGCTACCA TAACTTATCG TATTTATGCC GATTTGGATT ATATACAGAA CAATGATATT
CCGGTAAATA CTACTTCTGC TATCGGCCCG GATCTTGGTG GATTCGATAC CAATACCGAG
                                                                                        1500
                                                                                        1560
65
       GCAAAATTGA CCTATACCAA TTCCAATGGC GAACCGAATC AGCAGTTAAT TTTCCCACGT
       CCGACGGTTA AGTTAGGTTA TGGTGTTATT AAGCGGCACT ATGTATTGGT AAATAAAGAC
       GGTCAACCCA TACAGGCAAA TGGAACAGTT GTCAGTTCCC TAAGCGAGGC TCATGTTCTC CAGTCACAAA ATTCTTTTT GCCCTCAGGT GGAGGTCATA TTGTTCCCAA ATGGATAAAG TTGGACAAAA CGACCGAAGC ATTACAGTAC TATTCCGTAC CGCCGACTAA CACGGTCATC ACTACTGCCG ATGGTAAACG TTATCGTTTT GTCGAAGTCC CAGGCTCCAC GCCGAATCCG
                                                                                        1740
                                                                                        1800
70
       GGCCAAATCG GTATCAGTTG GAAAAAACCG GCAGGAAACG CTTACTTCGC TTACAAGCTC
                                                                                        1980
       CTCAATTATT GGATGGGAGG AACAACAGAC CAACAGAGTG AATGGGATGT GACGTCCAAT
                                                                                        2040
       TGGACAGGAG CCCAAGTACC GCTCACAGGA GAAGATGTAG AGTTTGCAAC GACAGAAAAT
                                                                                        2100
       TTCGGTTCTC CGGCGGTAGC CGATTTGCAT GTCCCGACAA CCAACCCCAA AATTATCGGT
                                                                                        2160
75
       AACCTTATCA ATAATTCCGA CAAGGATTTA GTTGTTACCA CAAGCAGTCA ATTGACGATC
                                                                                        2220
```

2280

```
AACGGCGTGG TTGAGGATAA CAATCCGAAT GTCGGTACGA TCGTCGTGAA GTCGTCGAAA
       GACAATCCTA CGGGGACATT GCTTTTTGCC AATCCGGGCT ATAATCAAAA TGTAGGGGGG
                                                                                   2340
       ACCGTCGAGT TTTACAATCA GGGATATGAT TGTGCCGATT GTGGTATGTA TCGCAGGAGC
                                                                                    2400
       TGGCAGTATT TCGGTATCCC TGTCAATGAA TCAGGTTTTC CAATTAATGA TGTGGGCGGA
                                                                                    2460
 5
       AACGAGACCG TCAACCAATG GGTTGAGCCT TTCAATGGCG ATAAGTGGCG GCCAGCACCT
                                                                                   2520
       TATGCACCTG ATACAGAGCT TCAAAAATTC AAGGGCTACC AGATCACGAA TGACGTGCAG
                                                                                    2580
       GCACAGCCTA CGGGAGTTTA CAGCTTCAAG GGTATGATTT GTGTGTGCGA TGCCTTCCTG
                                                                                    2640
      AATCTGACAC GCACGTCCGG TGTCAACTAC TCGGGCGCCA ACTTGATCGG CAACTCATAC
                                                                                    2700
      ACTGGAGCCA TCGACATCAA GCAGGGTATT GTCTTCCCGC CGGAAGTCGA GCAGACGGTG
                                                                                   2760
10
       TATCTGTTCA ACACGGGAAC ACGCGACCAG TGGCGTAAGC TTAATGGAAG CACGGTTTCA
                                                                                   2820
      GGCTATCGAG CCGGTCAGTA CCTCTCTGTA CCTAAGAATA CAGCGGGTCA GGACAATCTT
                                                                                   2880
      CCGGATCGTA TTCCATCGAT GCATTCCTTC TTGGTGAAGA TGCAGAACGG AGCGTCTTGT
                                                                                    2940
      ACGTTGCANA TCTTGTACGA TAAGCTGCTC AAGAACACGA CTGTAAACAA CGGTAATGGT
                                                                                    3000
      ACGCAGATCA CATGGCGATC CGGCAACTCC GGATCGGCGA ATATGCCGTC ACTTGTGATG GATGTTCTTG GTAACGAGTC GGCCGACCGT TTGTGGATCT TTACCGATGG GGGTCTTTCT
                                                                                    3060
15
                                                                                    3120
      TTCGGATTCG ACAACGGCTG GGATGGTCGC AAGCTGACTG AAAAAGGTTT GTCACAACTT
                                                                                   3180
      TATGCGATGT CTGACATCGG TAATGATAAA TTCCAGGTTG CAGGGGTTCC GGAGTTGAAT
                                                                                    3240
      AACCTGCTGA TCGGCTTCGA TGCGGATAAG GATGGTCAAT ACACGTTGGA GTTTGCTCTT
                                                                                    3300
      TCGGATCATT TTGCGAAAGG GGCTGTTTAC CTGCACGATC TTCAGTCAGG AGCCAAACAC CGTATTACGA ATTCTACGTC GTATTCATTC GATGCCAAGC GGGGAGATTC CGGGGCTCGT
                                                                                   3360
20
                                                                                   3420
      TTCCGCTTGT CATATGGATG TGATGAGAAC GTAGATGATT CGCATGTCGT GAGTACAAAT
                                                                                    3480
      GGCCGTGAAA TTATAATTCT GAATCAAGAT GCTCTTGACT GCACTGTAAC CTTATTCACA
                                                                                   3540
      ATAGAAGGTA AGCTTCTTCG CCGCTTGAAA GTATTAGCTG GTCATAGAGA AGTCATGAAA
                                                                                   3600
      GTGCAGACCG GAGGGGCCTA TATTGTGCAT CTTCAAAATG CTTTCACTAA TGATGTGCAT
                                                                                   3660
25
      AAGGTGCTTG TTGAGTAT
                                                                                    3678
       (2) INFORMATION FOR SEQ ID NO: 221
30
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 3675 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
35
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
40
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
           (im) FEATURE:
                  (A) MAME/KEY: misc feature
                  (B) LOCATION 1...3675
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:221
50
      ATGAAACGAT ATACAATAAT TCTTGCAGTT TTTCTTTTAT TCTGCACGGT ATTTACCTTT
                                                                                      60
      CAAATAAAAG CTCGCCCTTA TGAAAGATTT GCAGATGTAG AGAAGCCTTG GATTCAGAAA
                                                                                     120
      CATTCAATGG ATTCTAAATT GGTGCCTGCA AATAAGGGTA ACTTAATTCA AGCTGAAATT
                                                                                    180
      GTATACCAAT CTGTTTCTGA ACATAGTGAC TTAGTTATTT CACCTGTGAA CGAAATAAGG
                                                                                     240
55
       CCTGCAAATC GTTTCCCTTC GCATAGGAAG TCTTTTTTTG CAGAAAATCT ACGGGCATCT
                                                                                     300
       CCCCCGTAG TTCCCGTTGC CGTCGACAAG TATGCGGTAC CGGTTGCCAA TCCAATGGAT
                                                                                     360
       CCTGAAAATC CCAATGCCTG GGATGTGACG CTAAAAATCA CTACTAAAGC GGTAACAGTA
       CCTGTCGATG TGGTGATGGT TATCGACCAG TCTTCGTCAA TGGGAGGGCA AAACATTGCC
      AGATTAAAGT CTGCCATTGC ATCGGGACAG CGTTTTGTGA AAAAAATGTT GCCTAAGGGG
60
       ACGGCTACAG AAGGGGTGCG TATCGCTCTT GTGAGTTATG ACCATGAGCC TCATCGCTTA
                                                                                     600
       TCTGATTTTA CCAAAGACAC TGCTTTTCTC TGTCAAAAAA TCCGGGCTTT GACTCCTATT
                                                                                     660
       TGGGGAACAC ATACCCAGGG GGGGCTTAAA ATGGCGAGAA ACATTATGGC CACTTCTACT
                                                                                     720
      GCTGTGGATA AGCATATCAT ATTGATGTCT GACGGGTTAG CGACGGAGCA GTATCCTGTT
                                                                                     780
      AAAAATGTAA CTACTGCAGA CTTCATTGGC AAAACTGGAA ATGCGAATGA TCCCATTGAT
TTGGTTATAC AAGGAGCAAT TAATTTCCCT ACAAATTATG TTTCCAACAA TCCATCTACA
CCTCTTACCC CAAATTATCC AACTCATTCT TCTAAAGTTG GACGGAGAAA TCTGCCGGAA
TCCAAATTCG ATTATAGTAA TCTGAGTGCA AGGATTACTT TTGATGGTGT TGCTGGCGCA
                                                                                     840
65
                                                                                     900
                                                                                     960
                                                                                   1020
       TTGGTCTATG AACCGAGGTT TCCTCATCCC TATTATTATT ATTTCCCTTG TAACGCTGCT
                                                                                   1.080
      ATCAATGAGG CTCAGTTTGC GAAAAACTCT GGTTATACAA TCCATACTAT TGGCTATGAC
                                                                                   1140
70
       CTGGGAGATT TTGCCTTGGC CAACAATTCG TTGAAACTAA CCGCTACAGA CGAGAATCAC
                                                                                   1200
      TTCTTTACGG CGACACCGGC CAATTTAGCT GCAGCGTTTG ATAATATTGC CCAAACTATT
                                                                                   1260
       AATATAGGTA TACAGAGGGG GGAGGTGACG GACTTTGTAG CTCCTGGTTT CATCGTTAAA
                                                                                   1320
       AATCTGACGC AATCGGGAGA TGTTACTCAT TTGCTAAATG TTTCAAATGG AACGGTGCAC
                                                                                   1380
       TATGATGTCT CTACTAAAAA ACTGACATGG ACTACTGGTA CTATCCTGAG CTCATCAGAA
                                                                                   1440
75
       GCTACCATAA CTTATCGTAT TTATGCCGAT TTGGATTATA TACAGAACAA TGATATTCCG
                                                                                   1500
```

```
GTAAATACTA CTTCTGCTAT CGGCCCGGAT CTTGGTGGAT TCGATACCAA TACCGAGGCA
                                                                                   1560
        AAATTGACCT ATACCAATTC CAATGGCGAA CCGAATCAGC AGTTAATTTT CCCACGTCCG
                                                                                   1620
        ACGGTTAAGT TAGGTTATGG TGTTATTAAG CGGCACTATG TATTGGTAAA TAAAGACGGT
                                                                                   1680
        CAACCCATAC AGGCAAATGG AACAGTTGTC AGTTCCCTAA GCGAGGCTCA TGTTCTACAG
                                                                                   1740
        TCACAAGATT TCTTTTTGCC CTCAGGTGGA GGTCATATTG TTCCCAAATG GATAAAGTTG
  5
                                                                                   1800
        GACAAAACGA CCGAAGCATT ACAGTACTAT TCCGTACCGC CGACTAACAC GGTCATCACT
                                                                                   1860
        ACTGCCGATG GTAAACGTTA TCGTTTTGTC GAAGTCCCAG GCTCCACGCC GAATCCGGGC
                                                                                   1920
        CAAATCGGTA TCAGTTGGAA AAAACCGGCA GGAAACGCTT ACTTCGCTTA CAAGCTCCTC
                                                                                   1980
        AATTATTGGA TGGGAGGAAC AACAGACCAA CAGAGTGAAT GGGATGTGAC GTCCAATTGG
ACAGGAGCCC AAGTACCGCT CACAGGAGAA GATGTAGAGT TTGCAACGAC AGAAAATTTC
                                                                                   2040
 10
                                                                                   2100
        GGTTCTCCGG CGGTAGCCGA TTTGCATGTC CCGACAACCA ACCCCAAAAT TATCGGTAAC
                                                                                   2160
        CTTATCAATA ATTCCGACAA GGATTTAGTT GTTACCACAA GCAGTCAATT GACGATCAAC
                                                                                   2220
        GGCGTGGTTG AGGATAACAA TCCGAATGTC GGTACGATCG TCGTGAAGTC GTCGAAAGAC
                                                                                   2280
        AATCCTACGG GGACATTGCT TTTTGCCAAT CCGGGCTATA ATCAAAATGT AGGGGGGACC
                                                                                   2340
 15
        GTCGAGTTTT ACAATCAGGG ATATGATTGT GCCGATTGTG GTATGTATCG CAGGAGCTGG
                                                                                   2400
        CAGTATTTCG GTATCCCTGT CAATGAATCA GGTTTTCCAA TTAATGATGT GGGCGGAAAC
                                                                                   2460
       GAGACCGTCA ACCAATGGGT TGAGCCTTTC AATGGCGATA AGTGGCGGCC AGCACCTTAT
                                                                                   2520
       GCACCTGATA CAGAGCTTCA AAAATTCAAG GGCTACCAGA TCACGAATGA CGTGCAGGCA
                                                                                   2580
       CAGCCTACGG GAGTTTACAG CTTCAAGGGT ATGATTTGTG TGTGCGATGC CTTCCTGAAT CTGACACGCA CGTCCGGTGT CAACTACTCG GGCGCCAACT TGATCGGCAA CTCATACACT
                                                                                   2640
 20
                                                                                   2700
       GGAGCCATCG ACATCAAGCA GGGTATTGTC TTCCCGCCGG AAGTCGAGCA GACGGTGTAT
                                                                                   2760
       CTGTTCAACA CGGGAACACG CGACCAGTGG CGTAAGCTTA ATGGAAGCAC GGTTTCAGGC
                                                                                   2820
       TATCGAGCCG GTCAGTACCT CTCTGTACCT AAGAATACAG CGGGTCAGGA CAATCTTCCG
                                                                                   2880
       GATCGTATTC CATCGATGCA TTCCTTCTTG GTGAAGATGC AGAACGGAGC GTCTTGTACG
                                                                                   2940
25
       TTGCANATCT TGTACGATAA GCTGCTCAAG AACACGACTG TAAACAACGG TAATGGTACG
                                                                                   3000
       CAGATCACAT GGCGATCCGG CAACTCCGGA TCGGCGAATA TGCCGTCACT TGTGATGGAT
                                                                                   3060
       GTTCTTGGTA ACGAGTCGGC CGACCGTTTG TGGATCTTTA CCGATGGGGG TCTTTCTTTC
                                                                                   3120
       GGATTCGACA ACGGCTGGGA TGGTCGCAAG CTGACTGAAA AAGGTTTGTC ACAACTTTAT
GCGATGTCTG ACATCGGTAA TGATAAATTC CAGGTTGCAG GGGTTCCGGA GTTGAATAAC
                                                                                   3180
                                                                                   3240
30
       CTGCTGATCG GCTTCGATGC GGATAAGGAT GGTCAATACA CGTTGGAGTT TGCTCTTTCG
GATCATTTTG CGAAAGGGGC TGTTTACCTG CACGATCTTC AGTCAGGAGC CAAACACCGT
                                                                                   3300
                                                                                   3360
       ATTACGAATT CTACGTCGTA TTCATTCGAT GCCAAGCGGG GAGATTCCGG GGCTCGTTTC
                                                                                   3420
       CGCTTGTCAT ATGGATGTGA TGAGAACGTA GATGATTCGC ATGTCGTGAG TACAAATGGC
                                                                                   3480
       CGTGAAATTA TAATTCTGAA TCAAGATGCT CTTGACTGCA CTGTAACCTT ATTCACAATA
                                                                                   3540
35
       GAAGGTAAGC TTCTTCGCCG CTTGAAAGTA TTAGCTGGTC ATAGAGAAGT CATGAAAGTG
                                                                                   3600
       CAGACCGGAG GGGCCTATAT TGTGCATCTT CAAAATGCTT TCACTAATGA TGTGCATAAG
                                                                                   3660
       GTGCTTGTTG AGTAT
                                                                                  3675
40
       (2) INFORMATION FOR SEQ ID NO: 222
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1275 base pairs
                  (B) TYPE: nucleic acid
45
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
50
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
55
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1275
60
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:222
      ATGGAAGTGA AGAAAAACAC AGTGGTGCTA CGCCTTCTGA TTTGGTTCGT GGCCATTCTT
       CTCTTCCACT CCTCACGGCT GTGGGGACAG GAAGGGGAGG GGAGTGCCCG ATACAGATTC
                                                                                   120
65
       AAAGGATTCG TGGATACCTA CCATGCCGTA CGCAGCTCTT CTCCTTTTGA TTTCATGAGC
                                                                                   180
       TCGCGTACGA GAGTGAGAGG TGAGCTGGAG AGGTCGTTCG GTAATTCGAA AGTAGCCGTA
                                                                                   240
       TCGGTCAATG CCACCTACAA TGCTCTACTG AAAGACGAGA CCGGCTTACG TTTACGTGAA
       GCCTTCTTCG AGCATCAGGA AGAGCATTGG GGGTTGCGCC TCGGACGACA GATTGTCATT
                                                                                   360
      420
70
      TTTCTGGCAC AGGATTACGA TGATATTCGT ATGCCGGTCA ATGCATTGCG TTTCTCTGTC
                                                                                   480
      TTCAACGAAT CGATGAAAGT GGAAGTCGTG GTACTGCCTG TATTCGAGGG GTACCGTCTG
CCTGTGGATC CTCGCAATCC TTGGAATATC TTCTCCCTTT CGCCCATTGC TCAGGGGATG
                                                                                   540
                                                                                   600
      AATATCGTCT GGAAAGAAGA AGCCGGCAAA CCGGCCTTCA AGGTTGCCAA TATCGAGTAC
                                                                                   660
      GGTGCGCGAT GGAGCACTAC GCTCTCCGGT ATCGACTTCG CTTTGGCTGC ATTGCATACA
                                                                                   720
75
      TGGAACAAGA TGCCCGTCAT CGAAGTACAG GGCATTGTGC CGACGGAAAT CATCGTTAGC
```

```
CCTCGCTATT ATCGTATGGG ATTTGTCGGC GGCGACCTCT CCGTACCCGT CGGACAGTTT
                                                                                840
      GTTTTCAGGG GAGAGGCTGC GTTCAATATC GACAAACACT TCACCTATAA GAGTCATGCC
                                                                                900
      GAGCAAGAGG GTTTCCAAAC AATCAATTGG TTGGCCGGAG CCGATTGGTA TGCTCCCGGT
                                                                                960
      GAATGGATGA TCTCAGGACA ATTCTCAATG GAAAGCATAT TCAGGTATAG GGATTTCATC
                                                                               1020
      TCCCAAAGAC AACATTCTAC CCTGATTACT CTCAATGTTT CCAAGAAATT CTTCGGCAGT
                                                                               1080
 5
                                                                               1140
      ACACTCCAAC TTTCGGACTT CACCTACTAC GACCTTACGG GCAAAGGATG GTTCAGTCGC
      TTTGCAGCTG ACTATGCCTT GAACGATCAG ATACATCTGA TGGCCGGATA TGACTGGTTC
                                                                               1200
                                                                               1260
      AGTAGTAAGG GCAGCGGTAT ATTCGATCGC TACAAAGACA ATTCCGAACT CTGGTTCAAA
      GCCCGCTACA GCTTC
                                                                               1275
10
      (2) INFORMATION FOR SEQ ID NO: 223
           (i) SEQUENCE CHARACTERISTICS:
15
                 (A) LENGTH: 1212 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDHESS: double
                 (D) TOPOLOGY: circular
20
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: HO
          (iv) ANTI-SENSE: NO
25
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
30
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1\overline{2}12
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:223
35
      ATGTCCTCCT GTGAGGTGGC TTATTTTTCA CTAAAGCCGA TCGATCTGCA GAACATCCGC
      GAACGGAATC ACTCTTCCGA CATCGCGCTT TCCAATTTAT TAGACAATTC GAATCAGCTA
                                                                                120
      TTAGCTACTA TTCTGATCGG GAATAATGTG ATTAATGTAG CCATCGTTAT CCTTTCCAAT
                                                                                180
      TATGCCATCG AGCAGACATT CGTTTTCTCT TCTCCGATCA TTGGATTTCT GATCCAGACG
                                                                                240
      ATACTCCTGA CCACTGTTCT TTTGCTGTTC GGAGAGATTC TGCCGAAAGT GTATGCGCGG
                                                                                300
40
                                                                                360
      AAGAATCCGC TGCAATACTC GCGCTTTTCT GCTGCAGCTA TGTCCGTTAT CTATAAGATA
      TTGTCACCGT TTTCAAAATT GCTGGTCAAA AGTACCGGCA TCGTTACCAG AGGTATCAGC AAGAAGAAAT ACGATATGTC CGTGGATGAG CTCTCGAAAG CGGTAGCCCT CACCACTACG
                                                                                420
                                                                                540
      GAGGGAGAGC CGGAGGAGAA AGAAATGATT AACGAAATCA TCAAATTCTA TAATAAGACA
      GCCTGCGAAA TCATGGTTCC GCGTATCGAT ATTGTGGATG TGGATCTGAG CTGGCCATTT
                                                                                600
45
      CGTAAGATGC TTGACTTCGT TGTTTCGTCG GGTTATTCCA GACTTCCCGT TTCAGAGGGG
                                                                                660
                                                                                720
      TCAGAAGACA ATATCAAAGG GGTGATTTAC ATCAAAGATC TAATCCCACA CATGGATAAA
                                                                                780
      GGCGATGAAT TCGACTGGCA TCCTCTGATT CGTAAAGCAT ATTTTGTCCC CGAAAACAAG
      CGCATAGATG ATTTGCTCGA GGAGTTCAGA GCCAATAAGG TGCATGTCTC CATCGTTGTG
                                                                                840
      GATGAGTTCG GTGGCACTTG CGGACTGATC ACAATGGAGG ACATATTGGA AGAGATCGTC
                                                                                900
50
      GGCGAGATTA CGGACGAGTA CGATGAGGAA GAACTCCCCT TTAAGGTTTT GGGGGATGGC
                                                                                960
      AGTTATCTTT TCGAAGGAAA AACGTCTCTC TCCGATGTTC GACACTATCT TGACCTTCCG
                                                                               1020
      GAAAATGCTT TCGGTGAATT GGGGGACGAG GTAGATACGC TAAGTGGGCT CTTCTTGGAA
                                                                               1080
      ATCAAGCAGG AACTCCCCCA TGTGGGCGAT ACAGCAGTGT ACGAGCCATT CCGCTTTCAA
                                                                               1140
      GTGACCCAAA TGGACAAGCG CCGAATCATC GAAATCAAGA TTTTCCCTTT CGAGCGCACT
                                                                               1200
55
      TGGGAGGTCG AA
      (2) INFORMATION FOR SEQ ID NO: 224
60
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 780 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
65
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
70
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
75
           (ix) FEATURE:
```

WO 99/29870 PCT/AU98/01023

```
(A) NAME/KEY: misc feature
                   (B) LOCATION 1...780
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:224
  5
        ATGAAACTAT TACTTTATCT CCTATTGGTC TTGTCGACTC TATCCCCGAT GTATTCGCAA
                                                                                      60
        ATGCTCTTCT CAGAGAATCT CACAATGAAT ATAGACAGCA CCAAAACCAT ACAAGGAACG
                                                                                    120
        ATATTGCCCG TACTGGATTT CAAAACCGAA AAGGAAAATG TGTTCACCTT CAAAAATACT
                                                                                     180
        GCCAATCTCA ATCTGCTGAT AAAGCACGGT CAAGTAATCA ACTTAATTAA TAAGCTTGAG
                                                                                     240
 10
        TTTTCTACCT ATGGCAATAA AGTAACCGTA AGTGGAGGAT ATGTACACAC CGAATACCGC
                                                                                     300
        TATTTGTTGC ATCATGTTTT TGAGGTTTAT CCTTATGTCG AGTCGCAATG GGCAGAAAGT
                                                                                     360
        AGAGGAATGA AATATAAGGT TTCTACGGGA TTACAGTCGC GTTATCGGCT GGTAAATAGT
                                                                                     420
        GATAACTGTC TCATGTTTGC AACATTGGGG GTATTTTTCG AATTCGAAAA GTGGGAACAG
                                                                                     480
        CCAGCCACTA GCCTCTTTGC AGGAACGTAT GCATACAGCC GAAGTATCAA AAGCCACCTG
                                                                                     540
 15
        TCTATCAGTT TCAGACATCG GTTGGGTGAA CATTGGGAAT TTACAACTAC GGCTATTCAC
        CAGGGAAAGC CTGACAGTTA TTTTAAGAAG GCACGTTTTG GAGGAGCTAT CGACCTCAAA
TACCATATCA CACCTACGAT AGGAATACGC GGGGCCTATC GGATCATCTA CGATACTGCC
                                                                                    660
                                                                                    720
        CCTATTGTAC CTGTGCGGAA AGATTACAAC ACCGTTGATG TTGGTATCGA TATTTCGTTT
                                                                                    780
 20
        (2) INFORMATION FOR SEQ ID NO:225
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 2502 base pairs
 25
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
 30
           (iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
35
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
40
                  (B) LOCATION 1...2502
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:225
       ATGAAACGAA TCGTTTTATC ATCTTTCCTG TTCGTTCTGT CCATACTTTC TTTGATGGCA
45
       CAGAACAATA CCCTCGATGT ACACATATCC GGTACGATCA AGGATGCCTC CTCCGGCGAA
                                                                                    120
       CCAGTGCCCT ATGCCACTGT AAGCATCCGG CTGACAGGAG CAGATACCAC ACAGGTGTTC
                                                                                    180
       CGACAAGTGA CTGACGGCAA CGGCTACTTC GTCATAGGCC TGCCGGCAGC TCCCTCCTAT
                                                                                    240
       CACCTGACAG CTTCGTTCGT AGGTATGAAA ACCCATACCA TGCAGATTAG TCGGGGAAAT
                                                                                    300
       GGACAGCACG ACATCAAATC CATCGACATT TCTCTCGAAT CCGAGGACAA ACAACTCTCC ACCGTCACCG TATCGGCAGC ACGACCACTG GTGAAGATGG AGATAGACCG CCTGTCCTAT
                                                                                    360
50
                                                                                    420
       AATATGAAAG ATGACCCCGC AGCCAAGACG AACAACCTGC TCGAAATGCT GCGCAACGTT
                                                                                    480
       CCTTTGGTAA CGGTGGATGG TCAGGGCAAT ATCCAGGTGA AAGGATCTTC CAACTTCAAA
                                                                                    540
       ATCCACCTCA ATGGCAGGCC CTCGACCATG GTGAGCAGCA ACCCGAAGGA GGTCTTTCGC
                                                                                    600
       TCCATTCCTG CCCATACGAT CAAACGGGTG GAGGTCATCA CCGATCCGGG TGTAAAGTAC
                                                                                    660
55
       GATGCGGAAG GCACAAGTGC CATCCTGGAC ATCGTCACGG AAGAAGGTAA GAAGCTGGAA
                                                                                    720
       GGATATTCAG GTTCCATCAC GGCCAGTGTC AGCAACAATC CCACAGCCAA CGGTAGTATC
                                                                                    780
       TTTCTGACGG CAAAGTCCGG CAAAGTCGGG CTGACTACCA ACTATAACTA CTACGGTGGC
       AAAAACAAGG GCTCTCGCTA CTTTACCGAA CGTACTACAT CCATGCTCCA AACGATAGAA
                                                                                    900
       GAAGGCAAAG GGCAAGAAAC CTTTGGCGGA CACTTCGGCA ATGCCCTCCT CTCATTCGAG
                                                                                    960
       ATAGATTCGC TCAATCTCTT TACGGTGGGC GGCAATGTAC GCCTTTGGGA GATGACCACC
GACCGGAACA GCGTAGAAAA AAGCTTTGCC GGCAGCAACC TCATGTCCTA CATAGACAGA
60
                                                                                  1020
                                                                                  1080
       AAACTCAAAA CACAGATGGA TGCCGGATCA TACGAGCTCA ATGCCGACTA TCAGCACAGC ACTCGCCTGC CGGGCGAATT GCTCACCGTT TCCTACCGCT TCACTCACAA TCCTAATAAT
                                                                                  1140
                                                                                  1200
       AGCGAGACCT TCATTGACCA ATGGAAGCGC GATCCGCTCA ACACAGCTAA TACGATCCAG
                                                                                  1260
65
       TACGCCGGCC AGCACTCCAA ATCCGATGCG GGCATGGACG AACATACGGC ACAAGTGGAC
                                                                                  1320
       TATACACGTC CCTTAGGACA AGCACATTCT TTGGAAGCAG GGCTGAAGTA CATCTATCGT
                                                                                  1380
       CATGCCACGA GCGATCCTCT CTATGAGATA CGACCATCCG AAGATGCTCC GTGGCAGCCC
                                                                                  1440
       GGCTCTCTAT ATGCACAGAA TCCGTCGAAC GGAAAGTTCC GCCACGATCA ATACATCGGA
                                                                                  1500
       GCAGCCTATG CCGGCTACAA CTATCGTAAG GATCAGTATT CTTTGCAAAC CGGCCTCCGA
                                                                                  1560
70
       GTGGAAAGCA GCAGGCTGAA AGCACTCTTT CCCGAAAACG CAGCAGCAGA TTTCTCCCAC
                                                                                  1620
       AACTOGTTOG ACTGGGTGCC ACAGCTCACG CTCGGCTATA CCCCCTCGCC CATGAAGCAG
                                                                                  1680
       CTTAAGCTGG CCTATAACTT CCGAATCCAA CGTCCTGCAA TCGGCCAACT GAATCCCTAC
                                                                                  1740
       CGGCTACAGA CCAACGATTA TCAAGTACAG TATGGTAATC CCGACCTAAA GTCGGAGAAG
                                                                                  1800
       CGTCACCACG TCGGTCTCTC CTATAATCAA TACGGAGCCA AGGTCATGCT TACAGCATCG
                                                                                  1860
75
       CTCGACTACG ACTTCTGCAA CAACGCCATC CAGAATTACA CCTTCTCCGA CCCGGCCAAT
                                                                                  1920
```

```
CCCAATCTGT TCCACCAGAC CTATGGCAAT ATCGGACGAG AGCATTCTTT CAGCTTGAAT
      ACCTATGCCA TGTACACGCC GGCCGTATGG GTCAGGATTA TGCTCAACGG AAATATCGAT
      CGCACATTCC AAAAGAGCGA AGCACTCGGC ATTGATGTCA ATTCATGGTC CGGCATGGTA
                                                                                    2100
      TACTCAGGCC TGATGTTCAC CCTGCCGAAG GATTGGACTG TGAATCTCTT CGGAGGTTAT
                                                                                    2160
      TATCATGGGG GAAGAAGCTA CCAGACGAAG TATGATGGCA ATGTATTCAA CAATATCGGT
 5
                                                                                    2220
                                                                                    2280
      ATAGCCAAAC AGCTTTTCGA CAAAAAATTG AGAGTCTCGC TGAGCGCAAA CAACATTCAT
      GCGAAGTATT CGACATGGAA GAGCCGGACC ATCGGCAATG GATTTACTAT TTATTCGGAA
                                                                                    2340
      AATGCCGGTA TACAACGGAG TGTTTCCCTC AGCCTCACCT ACAGCTTCGG TAAGATGAAT
                                                                                    2400
      ACACAAGTGC GCAAGGTAGA GCGTACGATC GTCAACGACG ACCTCAAGCA AACCTCATCC
                                                                                    2460
10
      CAAGGACAGC AGGGTGGCGG ACAAGGAAAT CCTACCGGCA AT
                                                                                    2502
      (2) INFORMATION FOR SEQ ID NO: 226
15
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1197 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
20
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
25
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
30
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1197
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:226
35
      ATGAGACTCT CTGCCATTCT TATCGCTTTG ATTGTGATGC TGCCTGCTGT GCTTAGCGGG
                                                                                     120
      CAGCATTATT ATTCCATGGC GGGAGAGCGA CTGGAGACGG ACAGCATTCG TCCGAACGAA
      CTCTCGGCAT CGATCCGAAG TGCGCTTTTC TTTCGGAACA ATGAATACAA TGCACGTTCG
                                                                                     180
      GTCAAAGGTT ATACGTTGCC GGGTGCACGG GTTTCCGCTT TTGCCTCTTA CTCGCTGCCG
GCAGCACATG GTGTGAAGCT TTCGCTCGGA GTATCTACCC TGAACTACTG GGGGGCAAGT
                                                                                     240
40
                                                                                     300
      CGCTATCCGG CCGGTATCGC TTATTCCGAT TTACCTTATT GGACGGACTA TAACGACTAT GTACGCTTGC GTATCCTGC TTATGTACAG GCCATGCTGA AGCCGACGGC CACGACTGCT
                                                                                     360
                                                                                     420
      CTCATGCTGG GCAATATAGC CGGTGGTACG GCTCACGGAC TGATCGAACC GATCTACAAT CCTGAGTTGG ATTTGACGGC TGATCCTGAA GCCGGTGTGC AATTTCGGGG TGATTGGACA
                                                                                     480
                                                                                     540
      CGTTTCCGAA TGGATGTTTG GGTCAATTGG ATGAGCATGA TTTTCAAAAA TGACAATCAT CAGGAGTCGT TTGTCTTTGG CTTGTCCACT ACTTCGAAAT TGTTATCGGG TGAAGGCAAA
45
                                                                                     600
                                                                                     660
      TGGCGACTCG AACTGCCCTT GCAGGCTATT GCCACGCATC GCGGCGGGGA ATACAACTGG
                                                                                     720
      GCGCAGCAGG ATACCGTGCA TACATGGGTC AATGGAGCTG TCGGACTTAA GCTTTCGTAT
                                                                                     780
      CGCCCTCGTA CCGACAAACC CATGCAGATT TGGGGATCTG CTTATGGTGT GGCAGCCTTG
                                                                                     840
50
      TCAAGCGGAG GATACTTCCC TTACGAAAGA GGGTGGGGCG GTTATCTTTC TCTCGGAATG
                                                                                     900
      GACTTGGAGC ACTTCGCTTT TCGTACCGAC TATTGGTACG GCAGGCATTA CGTTTCTCCC
                                                                                     960
      TTTGCTGCAC CTTTCGCCAA TTCCCTGACG TATGACAAAC AGCCTCTTAC GAACGGTTGG
                                                                                    1020
      GGCGATTATA TTCGTCTCTA TGCCGACTAT TCGTGGCGGA TGGCACGAAG TGTTTCGTTG
                                                                                    1080
      GCGGCTGTTG CTCGGGTATG GTTCCAGCCT TCGGATCGTT TTGCGATGAG CCACGCCTTG
                                                                                    1140
55
      GAACTGACGA TGCGTATCGA TCCCAAATTC CCAATAGCTT TTCTGAAAGG CAATCAT
                                                                                    1197
      (2) INFORMATION FOR SEQ ID NO: 227
60
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1146 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
65
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
70
           (iv) AHTI-SEHSE: HO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
75
           (ix) FEATURE:
```

```
(A) HAME/KEY: misc feature
                    (B) LOCATION 1...1146
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:227
  5
        ATGAACAAAT CGCTATTATC ATTGGCATGC CTCATCCTGT GCGGTATGCC GGCCATCGCC
        CAACAGACAG GACCGGCCGA ACGCAGCGGC GAGCCTTCTC TGGCCGAACG TGTATTCGGT
                                                                                           120
        CTGGAGCAGA AGCAGAAAAA GCTGAAGGTG TACTTAGGCA TACAGTCGTT CTACGACCAG
                                                                                            180
        CCGCTTGTCG ATGACGAATC CCATATCGGA CACTTCAAGG TACAGGAGCT GCGGATGTCT
                                                                                            240
 10
        GCTCATGGCG AACTGAACCG CCACCTCAGC TTCGACTGGC GACAACGTCT CAACCGTGCC
                                                                                            300
        GCCGACGGCA CTTCGTTTGC CGACAATCTC TCCAATGCCA TCGACATCGC AGGTGTGGAC
TGGCACCCGA ACGACAAGGT GTCTTTCTTC TTCGGACGTC AGTACGCGCG TTTCGGAGGG
                                                                                           360
                                                                                           420
        ATAGAATACG ACATGAACCC CGTAGAGATC TACCAGTACA GCGACCTTGT GGATTACATG
        ACCTGCTATA CTTCGGGCGT GAACTTCGCA TGGAACTTCC ACCCGGAACA GCAGCTGCAG
CTACAGGTAC TCAATGCTTA CAACAACGC TTCGCCGAC GCTACCAGT GACACCCGAT
GTCGCTACCG CCACGAGCTA CCCGCTCCTC TACTCGGCAC AGTGGAACGG TACCCTCCTC
                                                                                           480
                                                                                           540
 15
                                                                                           600
                                                                                           660
        GGAGGAGCAC TGCATATGCG TTACGCCGTG TCGATGGCTC ATCAGGCCCA AGAGCGTAAT
                                                                                           720
        ATGTGGTACT TCACTGCGGG CAACCTGTTC AATCCGGGCA AACGGATCAA CGGATACCTC
                                                                                           780
        GACCTCACCT ACTCGATCGA GGGATTGGAC GACAAAGGCA TTATGACTGC TCGCTACGGC
                                                                                           840
20
        AAGGGCAAGA CCCTCACGGA CGTCAAGTAC TATGCTCTGG TATCGAAGTG GAACTTCCGC
                                                                                           900
        ATTTTCGATC AGGTCAATCT CTTCCTCAAA GGCATGTACG AGAACGGCTA TGCGCCTGCC
                                                                                           960
        CAATACGGCG AGAGCAGCCA CACGCGCCAC TCCTACGGCT ATATGGGAGG GGTGGAATAT
                                                                                          1020
        TACCCTACGG AGACCAACTT CCGTCTGTTC GTCACCTACA TAGGACGGCA TTACCGGTAC
                                                                                          1080
        AGTGCGACCG AGACGGAAAG CACCAATGCT CTTCGCGCCG GTCTGATCTA TCAGATACCT
                                                                                          1140
25
        TTCTTA
                                                                                          1146
        (2) INFORMATION FOR SEQ ID NO: 228
30
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 666 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
                    (D) TOPOLOGY: circular
35
             (ii) MOLECULE TYPE: DNA (genomic)
            (iii) HYPOTHETICAL: NO
40
             (iv) ANTI-SENSE: NO
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
45
             (ix) FEATURE:
                   (A) MAME/KEY: misc_feature
                    (B) LOCATION 1...666
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:228
50
        ATGAAACATT TGTTTAAGTC GACATTAGTA CTTCTTTGTG CTCTTTCTTT TTCCGGTACC
                                                                                            60
        TATACCTTTG CACAAGAAAA TAATACAGAA AAGTCACGAT TTGATTTTTC TGTTAGGCTG
                                                                                          120
        GGACAGGGAT ATATTGCAGG TTCAACTACC AACCTGATGT ATGGGTATAC ATCTGCTAAC
                                                                                          180
       GATAGACTTT TGTCTGGTGC AATTTATCTG GGCTTGACAC CAAGTAAGAA AGAAAATGCA
ACCGGCGTAG CATTTCGTTT CTTATCHCCC TCTCCGGGTT ATTATGTCGA TATATCCGGC
                                                                                          240
55
                                                                                          300
       AAAGAAAATA CCTTGAATTA TGCGTTTTAC GTTGTCGGAG CATATAATAG AATAGCCATT
CCTATACGCC CTATCAAAAA TTTTAATTTC ATCTTCTCTA CAGAAGTCGG AATGGCTTGG
                                                                                          420
       ATGAGTCGTC ATGAGCAAAT TTACAATTCT ACTTCGCAGA CTTGGGATAA GCAGCGCAAG
TCGAGGTCGG GACTGGATTT TGGTCTCGGG ATGCATCTGC AATHICCACAT TAATAAGACC
                                                                                          480
                                                                                          540
60
        GTTTACTTTA TGGCAGGAAC CGATCTTACG TCTTGCATGT TCGGAAAAAG GATCAATGAC
                                                                                          600
       TACCAGCAAA AGGATCGAAC CTTCATTGCA CTTATCGACA ACAGTATTGG CATAGGATTA
                                                                                          660
       AACCTC
                                                                                          666
65
       (2) INFORMATION FOR SEQ ID NO:229
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1173 base pairs
                   (B) TYPE: nucleic acid
70
                   (C) STRAHDEDNESS: double
                   (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
75
           (iii) HYPOTHETICAL: NO
```

```
(iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
 5
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1173
10
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229
      ATGATTATCA AGAAAATGCT GAAAAATAAA TTGGCCCCCT TGGCCATACT GTTCCTTTTT
      GCTCCAAAGG CTATGAAGGC TCAGGAGCAA CTGAATGTGG TACACACCTC TGTGCCATCG
                                                                                    120
15
      CTGAATATCA GTCCGGATGC ACGTGCGGCC GGTATGGGGG ATATAGGTGT GGCAACGACG
                                                                                    180
      CCGGATGCGT ATTCACAGTA TTGGAATCCG AGTAAATATG CTTTCATGGA TACGAAAGCC
                                                                                    240
      GGTATTAGCT TCTCATATAC ACCCTGGCTG TCCAAGCTGG TCAATGATAT TGCCCTGATG CAGATGACCG GTTTCTACAA ATTGGGAACA GACGAGAATC AGGCTATTAG TGCTTCTCTG
                                                                                    300
                                                                                    360
      CGTTATTTCA CATTAGGAAA GTTGGAGACT TTCGACGAAT TGGGCGAATC CATGGGAGAG
                                                                                    420
20
      GCCCATCCCA ATGAATTTGC TGTCGATTTG GGCTATAGCC GCCAGTTGTC GGAGAACTTC
                                                                                    480
      TCCATGGCTG TTGCACTGCG TTACATCCGC TCAGACCAAA GCACTCACAA CACCGGAGAG
                                                                                    540
      AATCAGGCCG GAAATGCCTT TGCGGCGGAT ATAGCCGGTT ATTTGCAGAA GTATGTGCTA
                                                                                    600
       CTGGGTAATG CGGAGAGCTT GTGGTCGTTG GGTTTCAACG TAAAGAATAT CGGAACGAAG
                                                                                    660
      ATCTCCTATG ACGGAGGTGT CACGAGTTTT TTCATCCCTA CTTCGTTGAA TCTCGGGACG
                                                                                    720
25
      GGGCTGTTGT ATCCGATCGA TGACTATAAC AGCATCAATT TCAACCTTGA ACTTAGCAAG
      CTGCTTGTAC CCACTCCTCC TATCATGGAT CAAAACGATC AGGCCGGGTA TGAGGCTGCA
CTCAAGAAAT ATCAGGAAAC TTCTTCGATC AGCGGTATAT TCTCTTCTTT CGGTGATGCG
                                                                                    840
                                                                                    900
      CCGGGAGGAC TCAAGGAAGA ATTCCGTGAG ATTACATGGG GACTTGGGGC TGAATATAGC TATGACGATA AATTTTTTGT TCGTGCCGGA TATTCATACC TGCACCCCAC CAAAGGCAAT
                                                                                    960
                                                                                   1020
30
      TTGCAGTACT TCACGGCCGG TGCCGGCTTC AAAATGAACA TATTCCGTAT CGATGCTTCC
                                                                                   1080
      TACCTGTTGT CTACGATCCA GAGTAATCCG TTGGATCAGA CTCTGCGGTT TACGCTTGCT
                                                                                   1140
      TTCGATATGG ATGGATTGCG CAATTTGTTC CAC
                                                                                   1173
35
       (2) INFORMATION FOR SEQ ID NO:230
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1338 base pairs
                  (B) TYPE: nucleic acid
40
                     STRANDEDNESS: double
                  (C)
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
45
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SEMSE: NO
           (vi) ORIGINAL SOURCE:
50
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1338
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230
      ATGAAAACAA CAGTTCAACA AATTATTCTG TGCCTGGCTT TAATGATGTC AGGTGTATTG
                                                                                     60
      GGCGGAAACG CACAGAGCTT TTGGGAAGAA ATAGCTCCTC CTTTTATCAG TAATGAGCCT
                                                                                    120
60
      AACGTCAAGT ATATAATTCC CAATATGGGG ATTGATTCAA AGGGAACAAT CTATGTAACC
                                                                                    180
      GTGACAAAAA GGATTCAGCA GGGAGCAAAT TATACTTCTG AGCAATTGGG TATGTACTAT
                                                                                    240
       CGACCATTAG GTGATAATGA ACAGTGGTGG AAACATGATC CGTATTTTGA TGACAAGATA
                                                                                    300
       GTTGCGGATA TTCAGACAGA TGCATATGGC AGAGTTTATG TATGTACGAC TTCTTCTCGA
                                                                                    360
      GATCAAGAGT ATCAACTTTA TATAAACGAG CAGAACGAAT GGAGGTGTAT ATTCAAAACT
65
      TCTGTGTCTA CATATGAGCA TGGTATGGCT GTTTTTCGCT CTTCGACAGG GGTGACTTAT
                                                                                    480
      ATAGGTACCA GGCATCACAT CTTCGCATCA GGTGTAAATG ATTTCGAGTT CAACACTATC
                                                                                    540
      TATGAAGACT CTACACCTAT GAGCTGTCGC TTTGCAGAGG CTACGAATAG TGGCACCATC TATCTGGCAT TGATGCATGA AACCACAATG TCTACGACTA TCCTTACTTA TCAAAACGGT
                                                                                    600
                                                                                    660
      GAGTTCGTCG ATATCTCGGA AAGTGAATTG AGTAACTCGA TTATTGCATC CATGTGCTCT
                                                                                    720
70
      AATAAAGAAG GTGATATAAT AGCTCTTGTT ACTTCATATA CAGGATTTAT GAGTGGAACC
                                                                                    780
      CTTGCGATCA GAAAAGCAGA TGAAGGCAAA TGGCAACTTG TTGGCGGAGA TATACAGAAT
                                                                                    840
      GCGATCGTTC AAAATATATG CATGATGGAC GACAACAAGA TTGCTTGTGA AGTCTTCGGG
                                                                                    900
       ACTCCTAACG GAGTAGATGG TCGGACAAGG GTTTGTGTTT CTGACGCATC TGTCTTTGAT
                                                                                    960
       TTTGAGTGGT ATGAAGATGA AATATACGGA GGCCTGATAT TTGACACTTT CTTCTATAGC
                                                                                   1020
75
       CCTTGGGACA AACTTCTTTA TGCGAAATTT GGTGGGATTA TGCTCAGGAG TAAAGAGTCT
                                                                                   1080
```

| 5 | TTTATAACCT CTTTCATTTC TCCGACAGTT GTACAAGGAG TGGATGTCTA TACTTTGGCC GGGAAGATAA GGATCGAAAG TGAAACTCCG GTGTCTGAGG TGTTGCTTTT CGACCTGGCT GGCAGGATGG TACTTCGGCA AACCATTGAT AATAAAATCT ATTCGGACAT AGATACTAAC GGACTAAAAGC GAAGCGGTAT TTACGTAGTC TCGGTGCGGC TCTCTTCCGG ACAGGTATTC AGTCATAAGG TGCAGGTA | 1140 1200 1260 1320 1338 | | | | | | | | |
|------------|---|--------------------------------------|--|--|--|--|--|--|--|--|
| | (2) INFORMATION FOR SEQ ID NO:231 | | | | | | | | | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 924 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | | | | | |
| 15 | (ii) MOLECULE TYPE: DNA (genomic) | | | | | | | | | |
| | (iii) HYPOTHETICAL: HO | | | | | | | | | |
| 20 | (iv) ANTI-SENSE: NO | | | | | | | | | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | | | | | | | | | |
| 25 | (ix) FEATURE: (A) NAME/KEY: misc feature (B) LOCATION 1924 | | | | | | | | | |
| 30 | (Mi) SEQUENCE DESCRIPTION: SEQ ID NO:231 | | | | | | | | | |
| | ATGATAATCC GGTGTCTTAT CCGTCGTCCG AGAACCGTCC TGTTCGGGTT GATATTCGTG GTAGGTCTTT TCTCTGCGAT GGCGCAAGAG AAAAAGGATA GTCTCTCTAC GGTTCAGCCA GTGCCGAATA GCAGCATGGT GGAGCAGACC CCTCTTCTCT CCATTGATCA CCCCGTCCTG | 60 120 180 | | | | | | | | |
| 35 | CCCGCTTCTT TTCAGAATAC CCGTACACTG AAAAGGTTTA GAGACAAACA TCTTTCCGAT GCTTTGCTCA ATGGATTGAA GCCTCATCGC TCATCTTTGC AATTGAATGA GGAACTCAAC TTCGCGGCAG AGCGTCGGGA TTTCGTTTCT CCCCTCTTGC AAACTCGCCA CGCTGCCGGT GTCCTTTCAT GGCGACCGAC CGATAGGATG CATTTTTATA CATCGGCCAA TATCGGTCTT | 240 300 360 420 | | | | | | | | |
| 40 | GGCCATGATT TATTGACCGG TGTGCGCAAG GACTTCGGAT GGAATGCTGG TGCCGACTTC TTGCTGAGTC AAAATCTTAC GGCACATGTC CAAGGCGGTT GGCAGCAGAA TTTCGGCTTT ATACCTATGA CGGCTGTCAA TGGCCAACTG CGTTGGCAAG CCACCGAGAG ATTGAGTTTT ACCACCGGTA TCGATTATCG ACAGGTACAG TGGAATGCTT TCGATAATAG AACGTTCTCG CTTAAAGGAA GTGCTCGATA CGAAGTGATG GACAATGTCT TTGTCAATGG ATTTGGCAGC | 480 540 600 660 720 | | | | | | | | |
| 4 5 | TATCCTCTCT ACAGCAGTAC GCGCTCAGGA CTCAATATGG CTGTTCCGAT GCATGGATTC GGCCCTCAGT ACGGTGGATC GCTTGAGCTG AAAGTCTCCG AGCGATTCGG CTTTGCCGTC GGTATGGAGC GCGAATACAA TATCTGGACT CGTCGGTGGG AAACGCATTA CTTTGCTTAT CCTGTATTCT ATGGCGATAA GAAG | 780 840 900 924 | | | | | | | | |
| 50 | (2) INFORMATION FOR SEQ ID NO:232 | | | | | | | | | |
| 00 | (i) SEQUENCE CHARACTERISTICS: (A) LEMGTH: 945 base pairs (B) TYPE: nucleic acid | | | | | | | | | |
| 55 | (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | | | | | |
| | (ii) MOLECULE TYPE: DNA (genomic) | | | | | | | | | |
| 60 | (iii) HYPOTHETICAL: NO | | | | | | | | | |
| | (iv) AHTI-SENSE: NO | | | | | | | | | |
| 65 | (VI) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROMONAS GINGIVALIS | | | | | | | | | |
| | (ix) FEATURE: (A) MANE/KEY: misc_feature (B) LOCATION 1945 | | | | | | | | | |
| 70 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:232 | | | | | | | | | |
| 75 | ATGANANCGA ATAGACGATA CGCATTTGTT TTGCCGCTTC TGCTACTCAC CGGATTGTTG GCATGGGGGC AGGATTCTTC CCACGGTAGC AATACAGCGT TTGCAACTGA TTCTTCGAGT AGAGAGTTGC CCACGGAGCA GTCCGCCTAC CGCATTCATT CTGCCTATAT GGTCGGTGGT GGCGGAAGCA TAACGCGCGA CACCTATTTG TCACCCCTTC GTTATGGAGG ATGGACACTG | 60 120 180 240 | | | | | | | | |

```
AATTTGTTGG GAGAGAAGAC GTTCCCTCTC AAAGCCTCCG ATTCCCGTTG GATGATCCGT
                                                                                     300
       ACCGGGCATG AGCTGGATTT TGCCCTGATG GACAATCCGG CCAATAATGC TCATTTCTAT
                                                                                     360
       TCCCTGCTGT ATAACGGTTC CGCTGCGGCT CTTTACCGCC TTGGCGCTAA GCATCTGCGA
       GCCGCGTGGA TGGACAATCT GCGCTTGGCA TTCGGCCCGG GCTTGGAAAT CGGGCTTGGA
 5
       GGAATTTATA GTACACGCAA CGGCAATAAT CCTGCGACAT TGAAGCTCTA CACCAATGCC
                                                                                     540
       ATCGCCCAAG CCTCGATAGG ATACTACGTC CCCTCCGAAA CTTTTCCCCT GTATTTTCGG
                                                                                     600
       TTGCTCTCCC AGATCAATCT CTTCGGTATA GCCTATGGAA ATGGTTTTGG TGAGAGCTAT
                                                                                     660
      TACGAGAATT TITTGCTCAA TAACGGCATT GCAGGCTCCC TGCATTTCAC TTATCCGGGC
AAGTTTACTC GGTTCACGAC ACTCATAACG GCGGATATTC CCATTCGGAA CTTCTGTACG
                                                                                     720
                                                                                     780
      CTTCGTGTCG GTTATCGCTA TTCCCATTTG GGCTCTTCGC TTAACGCATT GGATACTCGA ATCCACAGTC ATACGGCTT TATCGGTTTC GTCACGGAGT TTTACCGATT CCGTGGGCGC
10
                                                                                     840
                                                                                     900
       AAAGCCATGA ATACCGGTCG GAGAACCAGT CTTTACTATC ATGAT
                                                                                     945
15
       (2) INFORMATION FOR SEQ ID NO:233
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 855 base pairs
                  (B) TYPE: nucleic acid
20
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
25
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
30
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...855
35
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:233
      ATGGTAGTGA CGCTGCTCGT AATTGTCGGT ATTGTGGTAG TAGTGCGGTA CTCCTTGCGC
       GTCCACGTAC ATAAGACCGG CACTGTTGTA TCGGCGGCCA TATTTGGCTT CATCCTCCTT
                                                                                     120
40
       GGAAAGACCG TTCCATGCGA TACCCGTAAC TTCTTTTCCT CCGAAAGTGA TGAACCTGAG
                                                                                    180
       AGCCGTGTTG CTACCGAAAT AGCCCACCTG TGCGAAATAG GATTTCAGAT CCACGCTTCC
                                                                                     240
       TCTATCCACG TAGCCGTCCG AACCGATTTT GGACAGGCGG GCATCCACTG CCCAATGGCG
                                                                                     300
       ACCGATGCGT CCGCTACCGA GTTTGACCGA TCGGCGGAAT GTGCCGAACG AACCTCCGCT
                                                                                     360
       CAAATCGACA CGGCCATAAG GAGCCAGTCC CAAATTATCC GTACGCATAT TGACACTTGC
45
       CCCAAAAGCT CCGGCACCAT TGGTGGAAGT ACCCACACCT CGCTGCACCT GAAGGTCTTC
                                                                                     480
       GATGGAAGAG GCGAAGTCGG GCATATTCAC CCAAAAGACG GACTGAGATT CGGAGTCGTT
                                                                                     540
       GAGGGGTACT CCATTGGTAG TTATGTTGAT GCGATTGGCA TCGGTGCCAC GCACGCGAAA
                                                                                     600
      GCCGGAATAT CCGATACCCG TACCGGCATC GCTGGTGGCT ACCACGGAGG GAGTCAGCAT
                                                                                     660
      CAGCAGATAG GGGATGTCAC GACCATAATT GGACTTGGAA AGTTCGGCCT TGCGAACGTT
GGTGTAAGCG ACAGGGGTTT TCGCCGTGGC GCGAGTAGCT ACGACCTGTA CGGTCTGGAG
                                                                                     720
50
                                                                                     780
      CTGCACATTG CTAAGACTAT CTATCTCGCT GTTGGAGACG GGTGCTTGTG CCGTCAGGCA
                                                                                     840
      GAAAGGCAGG ACGGC
                                                                                    855
55
       (2) INFORMATION FOR SEQ ID NO: 234
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1797 base pairs
                  (B) TYPE: nucleic acid
60
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
65
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
70
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
(B) LOCATION 1...1797
75
```

```
ATGGAAAACT TAAAGAACAT TCAGCCCAGA GAGGATTTCA ACTGGGAAGA GTTTGAGGCC
                                                                             60
      GGTGGCGTCC ATGCTGCCGT GAGTCGTCAG GAGCAGGAAG CTGCTTATGA CAAAACGCTC
                                                                            120
 5
      AATACCATCA AGGAAAAGGA AGTGGTAATG GGTAGGGTAA CTGCTATCAA CAAGCGTGAA
                                                                            180
      GTGGTTATCA ATGTAGGGTA CAAATCGGAA GGTGTGGTAC CTGCAACAGA ATTCCGCTAC
                                                                            240
      AATCCCGAAC TCAAAGTGGG AGACGAAGTG GAAGTTTATA TCGAGAATCA GGAAGATAAG
                                                                            300
      AAGGGCCAGC TCGTCTTGTC TCACCGCAAG GGTCGTGCCG CTCGCTCTTG GGAGCGCGTG
                                                                            360
      AACGAGGCTC TCGAAAAAGA CGAAATCGTA AAGGGCTATG TGAAGTGTCG TACCAAGGGT
                                                                            420
10
      GGTATGATCG TCGATGTATT CGGTATCGAG GCTTTCCTCC CGGGATCACA GATCGACGTG
                                                                            480
      CGCCCCATTC GCGACTACGA TGCATTCGTT GAGAAGACGA TGGAGTTCAA GATTGTGAAA
      ATCAATCAAG AATATAAGAA TGTAGTTGTT TCCCACAAGG TGCTCATCGA AGCAGAGCTC
                                                                            600
      GAACAACAGA AGAAAGAAAT CATCGGCAAG CTCGAAAAAG GGCAGGTACT CGAAGGTATC
      720
15
      CATATCACTG ACCTTTCATG GGGTCGTGTG GCTCATCCGG AAGAAATCGT ACAGCTGGAT
      CAGAAGATCA ATGTCGTTAT CCTCGACTTT GATGAAGATC GCAAGCGTAT CGCTCTCGGA
                                                                            840
      CTCAAACAGC TGATGCCTCA TCCTTGGGAT GCTCTCGACA GCGAGCTTAA GGTAGGCGAT
                                                                            900
      AAGGTGAAGG GTAAAGTTGT GGTGATGGCA GATTACGGTG CTTTCGTTGA GATTGCACAG
                                                                            960
      GGCGTTGAGG GTCTTATCCA CGTAAGCGAA ATGTCATGGA CACAGCACTT GCGTTCTGCT
                                                                           1020
20
      CAGGACTTCC TGCATGTAGG CGACGAAGTG GAAGCCGTGA TCCTGACGCT CGACCGCGAA
                                                                           1080
      GAACGCAAAA TGTCGCTCGG TCTGAAGCAA CTCAAGCCGG ATCCTTGGGC TGATATCGAA
                                                                           1140
      ACTOGTTTCC CTGTAGGCTC TCGTCACCAT GCTCGTGTTC GCAACTTCAC CAATTTCGGT
                                                                           1200
      GTATTCGTTG AGATCGAAGA GGGCGTAGAT GGCCTTATCC ATATTTCCGA CCTTTCTTGG
                                                                           1260
      ACGAAGAAGA TCAAACACCC CAGCGAGTTT ACGGAAGTAG GTGCTGATAT CGAAGTTCAG
                                                                           1320
25
      GTAATCGAGA TCGACAAGGA AAACCGTCGT CTCAGCTTGG GTCACAAACA GTTGGAAGAG
                                                                           1380
      AATCCTTGGG ATGTATTCGA GACGGTATTC ACTGTAGGAT CTATCCACGA AGGAACGGTA
                                                                           1440
      ATCGAAGTGA TGGACAAGGG TGCTGTCGTT TCTCTGCCTT ACGGTGTGGA AGGTTTTGCC
                                                                           1500
      ACTCCGAAGC ACATGGTGAA GGAAGATGGC TCACAGGCTG TACTCGAAGA GAAGTTACCT
      TTCAAGGTTA TTGAGTTCAA TAAGGATGCC AAGCGAATCA TTGTATCTCA TAGCCGTGTA
                                                                           1620
30
      TTCGAAGATG AGCAGAAAAT GGCTCAGCGT GAAGCCAATG CAGAGCGTAA GGCTGAAGCC
                                                                           1680
      AAAGCGGCTC AGAAAGAAGC TGCTGCCGAA GCTGCCAATC CTGCACAGGC TGTAGAGAAA
                                                                           1740
      GCCACTCTCG GAGACCTCGG CGAGCTGGCC GCTTTGAAAG AAAAGCTTTC AGAAAAC
35
      (2) INFORMATION FOR SEQ ID NO:235
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1650 base pairs
                 (B) TYPE: nucleic acid
40
                    STRANDEDNESS: double
                 (C)
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
45
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
50
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1650
55
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:235
      ATGGGTAGGG TAACTGCTAT CAACAAGCGT GAAGTGGTTA TCAATGTAGG GTACAAATCG
      GAAGGTGTGG TACCTGCAAC AGAATTCCGC TACAATCCCG AACTCAAAGT GGGAGACGAA
                                                                            120
60
      GTGGAAGTTT ATATCGAGAA TCAGGAAGAT AAGAAGGGCC AGCTCGTCTT GTCTCACCGC
                                                                            180
      AAGGGTCGTG CCGCTCGCTC TTGGGAGCGC GTGAACGAGG CTCTCGAAAA AGACGAAATC
                                                                            240
      GTAAAGGGCT ATGTGAAGTG TCGTACCAAG GGTGGTATGA TCGTCGATGT ATTCGGTATC
                                                                            300
      GAGGETTICE TEEEGGATE ACAGATEGAE GTGCGCCCCA TTCGCGAETA CGATGCATTE
                                                                            360
      GTTGAGAAGA CGATGGAGTT CAAGATTGTG AAAATCAATC AAGAATATAA GAATGTAGTT
                                                                            420
65
      GTTTCCCACA AGGTGCTCAT CGAAGCAGAG CTCGAACAAC AGAAGAAAGA AATCATCGGC
                                                                            480
      AAGCTCGAAA AAGGGCAGGT ACTCGAAGGT ATCGTCAAGA ATATTACTTC TTACGGAGTA
                                                                            540
      TTTATCGACC TCGGTGGAGT GGATGGTCTT ATCCATATCA CTGACCTTTC ATGGGGTCGT
                                                                            600
      GTGGCTCATC CGGAAGAAT CGTACAGCTG GATCAGAAGA TCAATGTCGT TATCCTCGAC
                                                                            660
      TTTGATGAAG ATCGCAAGCG TATCGCTCTC GGACTCAAAC AGCTGATGCC TCATCCTTGG
                                                                            720
70
      GATGCTCTCG ACAGCGAGCT TAAGGTAGGC GATAAGGTGA AGGGTAAAGT TGTGGTGATG
                                                                            780
      GCAGATTACG GTGCTTTCGT TGAGATTGCA CAGGGCGTTG AGGGTCTTAT CCACGTAAGC
                                                                            840
      GAAATGTCAT GGACACAGCA CTTGCGTTCT GCTCAGGACT TCCTGCATGT AGGCGACGAA
                                                                            900
      GTGGAAGCCG TGATCCTGAC GCTCGACCGC GAAGAACGCA AAATGTCGCT CGGTCTGAAG
                                                                            960
      CAACTCAAGC CGGATCCTTG GGCTGATATC GAAACTCGTT TCCCTGTAGG CTCTCGTCAC
                                                                           1020
75
      CATGCTCGTG TTCGCAACTT CACCAATTTC GGTGTATTCG TTGAGATCGA AGAGGGCGTA
                                                                           1080
```

(xi) SEQUENCE DESCRIPTION: SEO ID NO:234

```
GATGGCCTTA TCCATATTTC CGACCTTTCT TGGACGAAGA AGATCAAACA CCCCAGCGAG
                                                                              1200
      TTTACGGAAG TAGGTGCTGA TATCGAAGTT CAGGTAATCG AGATCGACAA GGAAAACCGT
                                                                              1260
      CGTCTCAGCT TGGGTCACAA ACAGTTGGAA GAGAATCCTT GGGATGTATT CGAGACGGTA
      TTCACTGTAG GATCTATCCA CGAAGGAACG GTAATCGAAG TGATGGACAA GGGTGCTGTC
                                                                              1320
 5
      GTTTCTCTGC CTTACGGTGT GGAAGGTTTT GCCACTCCGA AGCACATGGT GAAGGAAGAT
                                                                              1380
      GGCTCACAGG CTGTACTCGA AGAGAAGTTA CCTTTCAAGG TTATTGAGTT CAATAAGGAT
                                                                              1440
      GCCAAGCGAA TCATTGTATC TCATAGCCGT GTATTCGAAG ATGAGCAGAA AATGGCTCAG
                                                                              1500
      CGTGAAGCCA ATGCAGAGCG TAAGGCTGAA GCCAAAGCGG CTCAGAAAGA AGCTGCTGCC
                                                                              1560
      GAAGCTGCCA ATCCTGCACA GGCTGTAGAG AAAGCCACTC TCGGAGACCT CGGCGAGCTG
                                                                              1620
10
      GCCGCTTTGA AAGAAAAGCT TTCAGAAAAC
                                                                              1650
      (2) INFORMATION FOR SEQ ID NO: 236
15
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1374 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
20
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
25
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORTPHYROHOMAS GINGIVALIS
30
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1374
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:236
35
      ATGATCGTCG ATGTATTCGG TATCGAGGCT TTCCTCCCGG GATCACAGAT CGACGTGCGC
      CCCATTCGCG ACTACGATGC ATTCGTTGAG AAGACGATGG AGTTCAAGAT TGTGAAAATC
                                                                               120
      AATCAAGAAT ATAAGAATGT AGTTGTTTCC CACAAGGTGC TCATCGAAGC AGAGCTCGAA
                                                                               180
      CAACAGAAGA AAGAAATCAT CGGCAAGCTC GAAAAAGGGC AGGTACTCGA AGGTATCGTC
                                                                               240
40
      AAGAATATTA CTTCTTACGG AGTATTTATC GACCTCGGTG GAGTGGATGG TCTTATCCAT
                                                                               300
                                                                               360
      ATCACTGACC TTTCATGGGG TCGTGTGGCT CATCCGGAAG AAATCGTACA GCTGGATCAG
      AAGATCAATG TCGTTATCCT CGACTTTGAT GAAGATCGCA AGCGTATCGC TCTCGGACTC
                                                                               420
      AAACAGCTGA TGCCTCATCC TTGGGATGCT CTCGACAGCG AGCTTAAGGT AGGCGATAAG
                                                                               480
      GTGAAGGGTA AAGTTGTGGT GATGGCAGAT TACGGTGCTT TCGTTGAGAT TGCACAGGGC
                                                                               540
45
      GTTGAGGGTC TTATCCACGT AAGCGAAATG TCATGGACAC AGCACTTGCG TTCTGCTCAG
                                                                               600
      GACTTCCTGC ATGTAGGCGA CGAAGTGGAA GCCGTGATCC TGACGCTCGA CCGCGAAGAA
                                                                               660
      CGCAAAATGT CGCTCGGTCT GAAGCAACTC AAGCCGGATC CTTGGGCTGA TATCGAAACT
                                                                               720
                                                                               780
      CGTTTCCCTG TAGGCTCTCG TCACCATGCT CGTGTTCGCA ACTTCACCAA TTTCGGTGTA
      TTCGTTGAGA TCGAAGAGGG CGTAGATGGC CTTATCCATA TTTCCGACCT TTCTTGGACG
                                                                               840
50
      AAGAAGATCA AACACCCCAG CGAGTTTACG GAAGTAGGTG CTGATATCGA AGTTCAGGTA
                                                                               900
      ATCGAGATCG ACAAGGAAAA CCGTCGTCTC AGCTTGGGTC ACAAACAGTT GGAAGAGAAT
                                                                               960
      CCTTGGGATG TATTCGAGAC GGTATTCACT GTAGGATCTA TCCACGAAGG AACGGTAATC GAAGTGATGG ACAAGGGTGC TGTCGTTTCT CTGCCTTACG GTGTGGAAGG TTTTGCCACT
                                                                              1020
                                                                              1080
      CCGAAGCACA TGGTGAAGGA AGATGGCTCA CAGGCTGTAC TCGAAGAGAA GTTACCTTTC
                                                                              1140
55
      AAGGTTATTG AGTTCAATAA GGATGCCAAG CGAATCATTG TATCTCATAG CCGTGTATTC
                                                                              1200
      GAAGATGAGC AGAAAATGGC TCAGCGTGAA GCCAATGCAG AGCGTAAGGC TGAAGCCAAA
                                                                              1260
      GCGGCTCAGA AAGAAGCTGC TGCCGAAGCT GCCAATCCTG CACAGGCTGT AGAGAAAGCC
                                                                              1320
      ACTCTCGGAG ACCTCGGCGA GCTGGCCGCT TTGAAAGAAA AGCTTTCAGA AAAC
                                                                              1374
60
      (2) INFORMATION FOR SEQ ID NO: 237
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1278 base pairs
65
                 (B) TYPE: nucleic acid
                 (C)
                    STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
70
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
75
           (vi) ORIGINAL SOURCE:
```

```
(A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
 5
                  (B) LOCATION 1...1278
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:237
      ATGGAGTTCA AGATTGTGAA AATCAATCAA GAATATAAGA ATGTAGTTGT TTCCCACAAG
10
       GTGCTCATCG AAGCAGAGCT CGAACAACAG AAGAAAGAAA TCATCGGCAA GCTCGAAAAA
                                                                                   120
      GGGCAGGTAC TCGAAGGTAT CGTCAAGAAT ATTACTTCTT ACGGAGTATT TATCGACCTC
                                                                                   180
       GGTGGAGTGG ATGGTCTTAT CCATATCACT GACCTTTCAT GGGGTCGTGT GGCTCATCCG
                                                                                   240
       GAAGAAATCG TACAGCTGGA TCAGAAGATC AATGTCGTTA TCCTCGACTT TGATGAAGAT
                                                                                   300
      CGCAAGCGTA TCGCTCTCGG ACTCAAACAG CTGATGCCTC ATCCTTGGGA TGCTCTCGAC
                                                                                   360
15
      AGCGAGCTTA AGGTAGGCGA TAAGGTGAAG GGTAAAGTTG TGGTGATGGC AGATTACGGT
                                                                                   420
      GCTTTCGTTG AGATTGCACA GGGCGTTGAG GGTCTTATCC ACGTAAGCGA AATGTCATGG
ACACAGCACT TGCGTTCTGC TCAGGACTTC CTGCATGTAG GCGACGAAGT GGAAGCCGTG
                                                                                   480
                                                                                   540
       ATCCTGACGC TCGACCGCGA AGAACGCAAA ATGTCGCTCG GTCTGAAGCA ACTCAAGCCG
                                                                                   600
       GATCCTTGGG CTGATATCGA AACTCGTTTC CCTGTAGGCT CTCGTCACCA TGCTCGTGTT
                                                                                   660
20
       CGCAACTTCA CCAATTTCGG TGTATTCGTT GAGATCGAAG AGGGCGTAGA TGGCCTTATC
                                                                                   720
       CATATTTCCG ACCTTTCTTG GACGAAGAAG ATCAAACACC CCAGCGAGTT TACGGAAGTA
                                                                                   780
       GGTGCTGATA TCGAAGTTCA GGTAATCGAG ATCGACAAGG AAAACCGTCG TCTCAGCTTG
                                                                                   840
       GGTCACAAAC AGTTGGAAGA GAATCCTTGG GATGTATTCG AGACGGTATT CACTGTAGGA
                                                                                   900
       TCTATCCACG AAGGAACGGT AATCGAAGTG ATGGACAAGG GTGCTGTCGT TTCTCTGCCT
                                                                                   960
25
      TACGGTGTGG AAGGTTTTGC CACTCCGAAG CACATGGTGA AGGAAGATGG CTCACAGGCT
GTACTCGAAG AGAAGTTACC TTTCAAGGTT ATTGAGTTCA ATAAGGATGC CAAGCGAATC
                                                                                  1020
                                                                                  1080
      ATTGTATCTC ATAGCCGTGT ATTCGAAGAT GAGCAGAAAA TGGCTCAGCG TGAAGCCAAT
GCAGAGCGTA AGGCTGAAGC CAAAGCGGCT CAGAAAGAAG CTGCTGCCGA AGCTGCCAAT
                                                                                  1140
                                                                                  1200
       CCTGCACAGG CTGTAGAGAA AGCCACTCTC GGAGACCTCG GCGAGCTGGC CGCTTTGAAA
                                                                                  1260
30
       GAAAAGCTTT CAGAAAAC
                                                                                  1278
       (2) INFORMATION FOR SEO ID NO: 238
35
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 720 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
40
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
45
           (iv) ANTI-SENSE: NO
           ("i) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
50
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...720
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:238
55
      ATGAAAAAAG CTATTCTTTC CGGAGCGGCC TTGCTCCTCG GCCTATGTGC CAACGCACAA
                                                                                    60
       AACGTGCAGT TGCACTACGA TTTCGGTCAT TCCATCTACG ACGAACTAGA TGGACGTCCC
                                                                                   120
       AAACTGACTA CCACAGTGGA AAACTTCACA CCCGACAAAT GGGGAAGCAC CTTCTTCTTC
                                                                                   180
       ATCGACATGG ATTACACGGG CAAGGGTATC CAGTCGGCCT ATTGGGAGAT TTCGCGCGAA
                                                                                   240
60
       CTGAAGTTTT GGCAAGCTCC CGTTTCCATT CATTTGGAGT ACAACGGAGG CCTCTCCACA
                                                                                   300
       AGCTTTACTT TOGGACACGA TGCTCTAATC GGTGCCACCT ACACCTACAA CAACCCCTCC
                                                                                   360
       TTTACACGTG GATTTACGAT CACGCCCATG TACAAGCATC TGGGTGCGCA CGACTTCCAC
                                                                                   420
      ACCTATCAGA TCACCGGCAC TTGGTACATG CACTTTCTGG ACGGTCTGCT TACCTTCAAC
                                                                                   480
       GGCTTCCTCG ATCTTTGGGG TTTCCCCCAA GAGAACCCAA TCGGGGGCCC TGTGCTCAAA
                                                                                   540
65
       GAAGGGGATA AGTTCGTATT CCTGTCCGAA CCGCAGTTCT GGATCAACCT CAATCGCATC
                                                                                   600
       AAAGGCATCG ACAAGGATTT CAATCTCAGC ATAGGGACAG AGATGGAAAT CAGCAGGAAC
                                                                                   660
       TTCGCTCGCA TGGACAAATT CTCCTGCATC CCTACTCTTG CGGTCAAATG GACTTTCAAC
                                                                                   720
70
       (2) INFORMATION FOR SEQ ID NO: 239
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1302 base pairs
                  (B) TYPE: nucleic acid
75
                  (C) STRANDEDNESS: double
```

```
(D) TOPOLOGY: circular
           (ii) HOLECULE TYPE: DHA (genomic)
 5
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
10
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1302
15
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:239
      ATGTATAAAG ACTATAAGGG TTTGTATGCG TCGCTTCGGT GGTATGCCCT GATCATTGGG
      TTGCTATTTG CAGCAGACGG TATACAGGCT CAGAACAACA ACTTTACCGA GTCGCCTTAC ACTCGCTTCG GCCTTGGCCG TCTCGGAGAA CGGACGACTA TTAGTGGGCA TTCCATGGGA
                                                                                    120
20
                                                                                    180
       GGACTCGGCG TCGGTCTGCG TCAGGGCACA TACGTCAATG CCGTCAATCC TGCTTCATAC
                                                                                    240
       TCGGCTGTGG ATTCGATGAC GTTTATCTTC GATTTCGGTG CATCTACCGG AATTACGTGG
                                                                                    300
       TATGCCGAGA ACGGGAAAAA GGACAATAGG AAAATGGGAA ACATTGAGTA TTTCGCCATG
                                                                                    360
       CTTTTTCCTA TTTCCAAATC CATTGCTATG AGTGCGGGAG TGCTTCCTTA CTCCGCATCC
                                                                                    420
25
       GGGTACCAGT TCGGATCCGT TGATCAAGTG GAAGGAGGCA GCGTCCAGTA CACCCGTAAA
                                                                                    480
      TACTTGGGGA CAGGCAATCT GAACGATCTC TATGTCGGTA TAGGTGCAAC CCCGTTCAAA
                                                                                    540
       AACTTCTCAA TAGGAGCCAA TGCTTCATCC CTTTTTGGGC GATTCACACA CAGCAGGCAG
                                                                                    600
      GTAATCTTCT CCACGGAGGC TCCTTACAAT CCCGTACATC TCTCGACGCT GTACTTGAAG
GCTGCCAAGT TCGACTTCGG TATGCAGTAT CACCTTCTTC TCAAATCAGA TCGTTCGCTC
                                                                                    660
                                                                                    720
30
       GTTATCGGTG CCGTCTATTC TCCGCGGGTG AAGATGCATA GCGAGCTGAC TCAGATAAAG
                                                                                    780
      AATCAGGTTC AGAACGGTGT AGTAGTGGAG AGCGAAACCC AAGAATATAT CAAGGGAATG
                                                                                    840
      GACTATTATA CCCTGCCTCA TACATTGGGG ATAGGTTTTT CTTATGAAAA GAAAGATAAA
                                                                                    900
      CTTCTCTTAG GAGCAGACGT CCAATATAGT AAATGGAAAG GCGAGAAATT TTATAAATCC
                                                                                    960
      GATTGCAAAT TCCAGGACAG AATACGGGTA TCTCTCGGCG GAGAGATCAT ACCGGATATA AATGCCGTTG GGATGTGGCC TAAAGTTCGC TATCGCTTCG GTTTACATGG TGAAAATTCT
                                                                                  1020
35
                                                                                  1080
       TACCTGAAAG TGCCGACTAA AGGCGGTGTA TATCAAGGAT ACCATATCGT AGGTGCTGTA
                                                                                  1140
       TTCGGTATAG GAATCCCGCT CAATGACAGA CGTTCGTTCG TAAATGTCTC TCTTGAATAT
                                                                                  1200
      GACCGATTGA TCCCGAAGGA GGGTATGATC AAAGAAAATG CTCTGAAATT GACCTTCGGC
                                                                                  1260
      CTCACGTTCA ACGAGTCATG GTTTAAAAAG CTGAAACTGA AC
                                                                                  1302
40
       (2) INFORMATION FOR SEQ ID NO:240
            (i) SEQUENCE CHARACTERISTICS:
45
                  (A) LENGTH: 2778 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDHESS: double
                  (D) TOPOLOGY: circular
50
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SENSE: NO
55
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
60
                 (A) HAME/KEY: misc_feature(B) LOCATION 1...2778
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:240
65
      ATGCGATCGA TTTATCAATT ACTGTTGTCA ATACTCCTTG CTTCTCTTGG TTTCGTCGGG
                                                                                    60
      CTGGAAGCCC AACAAGCCGG AGTAGCAGGT AGAGTATTGG ACGAAGAAGG CAACCCCATG
                                                                                   120
      ATTCAAGCCA ACGTACAGCT TGTACAGAGT ACCGGCCAAG TAGCCGTTGC CGCAGGTGCC
                                                                                   180
      ACTAATGAAA AAGGGTTGTT CAGCCTGAAA ACGTCACAGG AGGGTGACTA CATTCTGCGC
                                                                                    240
      GTTTCATATG TAGGTTACAC TACCCACGAC GAAAAAATAT CTCTTAGAAA CGGGCAAACC
70
      ATTACGCTCA AAGATATATC CATGAACGAA GATGCCCGTC TTCTACAGAG TGTGACGGTG
                                                                                   360
      CAGGCTAAAG CGGCAGAGGT CGTGGTACGC AACGATACGC TCGAATTCAA TGCCGGATCC
                                                                                   420
      TATACCGTAG CACAGGGAGC TTCTATCGAG GAACTGATCA AGAAGCTACC CGGAGCAGAG
                                                                                    480
      ATCGGATCCG ATGGGAAGAT CACCATCAAC GGCAAGGACA TTAGCAAGAT CCTTGTCGAT
                                                                                   540
      GGCAAAGAGT TTTTCTCCAA AGATCCACAG GTGGCAATAA AGAATCTTCC GGCCGATATG
                                                                                    600
75
      GTCAATAAAG TACAGGTACT GAACAAACTG AGCGAGCTGT CGCGGATGAG CGGTTTCGAT
                                                                                    660
```

```
TTCGGAACGC TTCAGGCCGG CTACGGTACC GACCAACGCT ATATGGCCGG AGGGAACGTC
                                                                               780
      AATCGGTTCG ATGGAAATAA GCAATGGACA TTGATCGGTA GTGCGAACAA TACGAACAAT
                                                                               840
      ATGGGCTTTA GCGAGATGGA CAGCGAGATG GGATCCATGA CCTTCTTCTC TCCCCAAGGC
                                                                               900
 5
      GGTGGTCGAC GCGGCTTCGG CAATAGTGGA GGTGTTACGT CTTCGTCGAT GCTGGGCGGC
                                                                               960
      AACTTCAGTG TCGAATTCTC CTCTGCCCTT AATACAGGAG GCGATGCACG CTACGGATAC
                                                                              1020
      AACGACAAGG CCATAGAGAC GACCAAACGC GTGGAAAATA TCCTCGCCGA AGGGAATACT
                                                                              1080
      TATATGGACG AAAATATATT GGAACGCTCT TTCTCTCACA ATGGTCAGGC GCGATTTAGG
                                                                              1140
      ATGCAATGGA AACCGTCCGA ACGTACCGAA GTGGTATTCG AGCCGGATCT TTCGATATCC
                                                                              1200
10
      AAGATCGATG GGTTCTTTAA CGACACATAC GAGACGAAAG ATGCCACCGG AATCTCTATC
                                                                              1260
      AACAAAGGTT CTATCCACCA AACTACACAA GGAAACAACT TCAGACTGAA CGGAGAATTG
                                                                              1320
      GATATCAGTC ACAAGCTCAA CGACGAAGGC CGTACGATCA GTGCCTCCGT CAGTGGCGGT
                                                                              1380
      CTGACCGACG AAGACGGAGA TGGCATATAT CAGGCTGTGC TCCAAAGCGT GGAGACGAAT
                                                                              1440
      CAAAAGCAAT TCAACGACAA CTCCAACCTG CAATATCGGC TTCGCCTCTC GTATGTGGAA
                                                                              1500
15
      CCGTTGGGTA AAAACTACTT CGCACAAGCG ATTCTGAACA GACGTTTCTC CCGTCGCAAT
                                                                              1560
      TCGGATCGTG AGGTGTACCG ACTGGGCGAT GACGGGCAAT ACTCCATATT AGACAGTCAG
                                                                              1620
      TACGGACTCT CCTACAGTAA CGAGTTCACC CAGTATCGCA TCGGACTCAA CCTCAAGAAG
                                                                              1680
      ATTGCCAAAA CGTGGGACTA CACCGTAGGA TTCAATGTGG ATCCCAACAG AACTGTCAGC
                                                                              1740
      TATCGGAGCG TAGCCGGAGT AGAGCAGGAC AAACTGGCTT TCAATCGTGT CAATCTCTCC
                                                                              1800
20
      CCGATGCTCC GAATCAACTA CAAACCGAGC AGGACTACCA ACCTCCGAGT GGACTACCGA
                                                                              1860
      GGACGCACGA CACAACCATC CATCAATCAG ATCGCTCCCG TTCAGGACAT CACGAATCCG
                                                                              1920
      CTATTCGTGA CGGAAGGCAA TCCCGGTCTG AAGCCGAGCT ATTCCAACAA TGTGATGGCC
                                                                              1980
      ATGTTCTCGG ACTTCGATGC CAAAAGTCAG CGAGCTTTCA ACATTGTTTT CTTCGGCAAC
                                                                              2040
      TATACATTCG ACGACATCGT CCCCAATACG CACTACGATC CGTCTACAGG GATCCGTACC
                                                                              2100
25
      ACTOSTRACS AAAACGCCTC CGGTACGTGG CAAGCGAATC TTCATGGGAC ACTATCGCTT
                                                                              2160
      CCACTCAAGA ACAGGGCATT TTCTTTCAGG ATGTCCTTGT TCAACAGGTT GGCCGAAGGA
                                                                              2220
      CAAAGCTTCA TCAATGACGA TAAGAACAAA GCTCTCTCTT TCCGAACGAG GGAACGCCTG
                                                                              2280
      ACGCTGACCT ATCGCAACAA TTGGATCGAT ACGAGTATCG GTGGCAATAT CGGATTCTAT
                                                                              2340
      ATGGCGAATA ATAGTCTGAG CGGACAGAAA GATTCTCGCA CATACGATTT TGGCGGCAAT
                                                                              2400
30
      TATCAAGTTG CCCTAACGCT TCCCTATGGA TTCCGTATCG ACAGCGATGT TGAATACAAT
                                                                              2460
      ACGAACTCCG GTTACAGCGG AGGATTCAGT CTGGACGAAT GGCTTTGGAA TGCTTCGCTT
                                                                              2520
      TCATACAGCT TCCTCCGTGA CAAGGCCGGT ACACTGCGTG TCAATGGCTA TGACATCCTC
                                                                              2580
      GGTCAGCGGT CAAGTATCAG CCGTTCTGCT TCGGCCATCA ATATAGAAGA GAGCATGTCC
                                                                              2640
      AATACGATCG GACGCTACGT GATGGTGGAC TTTATCTACC GATTCAACGC CTTCAGTGGT
                                                                              2700
35
      GGTGGATCTC GCAGCGATCA TCAGCGTGGC AATATGAATC GTCCGGGCCC ACCTTTCGGC
                                                                              2760
      GGTGGCAGAC GACCGTCC
                                                                              2778
      (2) INFORMATION FOR SEQ ID NO: 241
40
            (i) SEOUENCE CHARACTERISTICS:
                (A) LENGTH: 1200 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
45
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
50
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
55
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1200
60
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:241
      ATGGTTATGA AGCTGATTAA AAGAAGTTTG CTCCTGCTTG GAGCGGTACT GCTGATTACG
                                                                                60
      CTTCCTGCGT ACTCGCAGAA TGATGACATC TTCGAAGATG ACATCTATAC ATCGCGAAAA
                                                                               120
      GAAATACGTA AACAAAACCA AGTTAAAGAC TGGCAAAACC AAGAGGACGG ATACGGCGAC
                                                                               180
65
      GATACGGAAT ATACAGTGGC TTCCGATCGG GACATTGACG CCTACAATCG TAGAGATGGC
                                                                               240
      CAGTCCTACG ATGGGAAAAA GTTGTCCAAA GACAAGAAAA GAGACTCCAC TCGTTCTTCT
                                                                               300
      GTTCCCGGTC GCTATAGTCG CCGCTTGGCT CGATTCTATA AGCCGAATAC GATCGTCATT
                                                                               360
      TCAGGTGCCG ACAATGTATA TGTAACTGAT GATGGTGAGT ATTTCGTCTA TGGAGACGAA TACTATGATG ACGCGTCGTC TGTAAACATT TACATCAACA GTCCTTGGTG CGATCCGTTC
                                                                               420
                                                                               480
70
      CCTTATACGT CATGGTATCC ATCTTTCTCC GGCTGGTACA ACTATACGTG GAACTATCCA
                                                                               540
      TGGTTCTACT ACGGTAGCCA TATCGGATGG GGCGGTTATT ACCCCGGATA TAATTGGTAT
                                                                               600
      TGGAGCTACT ACTATGATCC TTTCTACAAT CCCTATGGAA TCGGTATGGG TTGGGGATAT
                                                                               660
      CCTTATGGCT GGGGCAGCTA TTACGGTTGG GGTGGCTATC CGGGAGTGAT ACATCACTAC
                                                                               720
      CACCACTACC CCAAGAAGAC CTATTCCAAT GGTCAGCATT CCGGAGCTTA CTATTCTTAT
                                                                               780
75
      GGCCGACCGA ATCGTATCAA AGGTGGAACG TCCGGTGCCA AACTTGGGAC AGGACGCTAC
                                                                               840
```

GATGGAGAG AGGAGACCGT AATCAACCTG ACGGTGAAGC CCGAAAAAAA GAAAGGCCTC

```
GATAGAATTC AAAATTCGTC TTCGCAAAAA AATAAGTTCG GATTGCAGTC GAACAAACCC
                                                                                         900
       AATAATAATC TGCAAAATGT CAAGTCGGGA CGTACCGGCC GAGCCAATAG AGACCGAAAT
                                                                                         960
       ATAGAAACGG TAACTCCAAA CAACGGGCAA AAGCAGAATC GTCCCGTATT CCAGCAGAAT
                                                                                       1020
       CAGTCCGGCA ATGACCGACC GACCGGACGG AATATCCGCA GCGAGAGACA GGGGGAAAAT
AACGATAGGA CATTTCGAC TCCTTCTCGT AGCAATAGTA ACGGTGGCTT CTCCACCGCCT
                                                                                       1080
 5
                                                                                       1140
       TCTCGCTCTT CTTCCGGCTC TATGAGCGGA GGTGGCGGAC GTAGTGGCCG GGGACGCAAT
                                                                                       1200
       (2) INFORMATION FOR SEQ ID NO: 242
10
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1194 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
15
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
20
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
25
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1194
30
           (xi) SEQUENCE DESCRIPTION: SEO ID NO: 242
       ATGAAGCTGA TTAAAAGAAG TTTGCTCCTG CTTGGAGCGG TACTGCTGAT TACGCTTCCT
                                                                                         60
       GCGTACTCGC AGAATGATGA CATCTTCGAA GATGACATCT ATACATCGCG AAAAGAAATA
                                                                                        120
       CGTAAACAAA ACCAAGTTAA AGACTGGCAA AACCAAGAGG ACGGATACGG CGACGATACG
                                                                                        180
35
       GAATATACAG TGGCTTCCGA TCGGGACATT GACGCCTACA ATCGTAGAGA TGGCCAGTCC
                                                                                        240
       TACGATGGGA AAAAGTTGTC CAAAGACAAG AAAAGAGACT CCACTCGTTC TTCTGTTCCC
GGTCGCTATA GTCGCCGCTT GGCTCGATTC TATAAGCCGA ATACGATCGT CATTTCAGGT
                                                                                        300
                                                                                        360
       GCCGACAATG TATATGTAAC TGATGATGGT GAGTATTTCG TCTATGGAGA CGAATACTAT
                                                                                         420
       GATGACGCGT CGTCTGTAAA CATTTACATC AACAGTCCTT GGTGCGATCC GTTCCCTTAT
                                                                                         480
40
       ACGTCATGGT ATCCATCTTT CTCCGGCTGG TACAACTATA CGTGGAACTA TCCATGGTTC
                                                                                         540
       TACTACGGTA GCCATATCGG ATGGGGCGGT TATTACCCCG GATATAATTG GTATTGGAGC TACTACTATG ATCCTTCTA CAATCCCTAT GGAATCGGTA TGGGTTGGGG ATATCCTTAT
                                                                                         600
                                                                                         660
       GGCTGGGGCA GCTATTACGG TTGGGGTGGC TATCCGGGAG TGATACATCA CTACCACCAC TACCCCAAGA AGACCTATTC CAATGGTCAG CATTCCGGAG CTTACTATTC TTATGGCCGA
                                                                                         720
                                                                                        780
       CCGAATCGTA TCAAAGGTGG AACGTCCGGT GCCAAACTTG GGACAGGACG CTACGATAGA ATTCAAAATT CGTCTTCGCA AAAAAATAAG TTCGGATTGC AGTCGAACAA ACCCAATAAT
45
                                                                                        840
                                                                                        900
       AATCTGCAAA ATGTCAAGTC GGGACGTACC GGCCGAGCCA ATAGAGACCG AAATATAGAA
                                                                                        960
       ACGGTAACTC CAAACAACGG GCAAAAGCAG AATCGTCCCG TATTCCAGCA GAATCAGTCC
                                                                                       1020
       GGCAATGACC GACCGACCGG ACGGAATATC CGCAGCGAGA SACAGGGGGA AAATAACGAT
                                                                                       1080
50
       AGGACATTTT CGACTCCTTC TCGTAGCAAT AGTAACGGTG GCTTCTCCAC GCCTTCTCGC
                                                                                       1140
       TCTTCTTCCG GCTCTATGAG CGGAGGTGGC GGACGTAGTG GCCGGGGACG CAAT
                                                                                       1194
       (2) INFORMATION FOR SEQ ID NO: 243
55
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1743 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
60
                  (D) TOPOLOGY: circular
            (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
65
           (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
70
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1743
75
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:243
```

```
ATGATCCGCA AGTTGATCTT ACTGCTTGCT CTGATGCCGG TAGCCTCTGT GGCTTTTGCG
      GTACCAACGG ACAGCACGGA ATCGAAAGAC AATCGTATCC TTACAAGCAT GCAATCCTCC
                                                                             120
      TCTTTGAATA GGGATGATGC TCCGGATAAA TGGCAACCTA TGCATGCCAA TTTCAGTATT
                                                                             180
 5
      CAGAGCGATA TGCTGCTTTC TACTGCCCAA AAGTCCAAGA ACACCTGGTT CGGCAACTCC
                                                                              240
      TATATCATGG GTATAATCAA GAACAATTAT CTGGAGTTTG GTGCCCGTTT CGAGGATCTC
                                                                              300
      TATAAGCCCC TGCCCGGACA TGAACCCGAG ATGGGGCGTG GCGTTCCTCA CATGTATGTG
      AAGGGAAGCT ATCATTGGGC GGAGCTGACT ATGGGAGACT TCTACGATCA GTTCGGTAGC
                                                                              420
      GGTATGGTAT TCCGCACCTA TGAAGAGCGC AACCTCGGTA TAGACAACGC GGTTCGCGGC
10
      GGACGTATAG TACTCACTCC TTTTGATGGA GTGCGTGTCA AGGGTATTGC AGGACAGCAG
      CGTAACTACT TCGACCGCAC GGGCAAGGTA TTCAATTCCG GCCGAGGCTA CCTACTGGGT
      TCTGATCTGG AGCTGAATGT AGAGCGTTGG AGCAGTGCCA TGCGCGACAA TGACTATCAT
                                                                             660
      TTGGCTATCG GGGGATCGTT CGTTTCCAAA CACGAAGCAG ACGAAGATAT ATTTGTGGGT
      GTAGGCGAAG ATCGCAAGCG ACTCAACCTG CCGCTCAATG TCCCGATTAT GGGCCTGCGC
                                                                              780
15
      ACCAACTITC AAAAAGGAGG TCTCGCCCTC TACGCAGAGT ATGGATACAA ATACAACGAT
                                                                             840
      CCCTCGGCAG ACAATGACTA TATCTACCAC GACGGACAGG CTGCACTCCT CTCTGCCTCA
                                                                             900
      TACTCCAAAA AAGGGATGAG TATCCTGTTG CAGGCCAAAC GTTGTGAGAA CTTTGCTTTC
                                                                             960
      CGCAGCAAGC GAAGTGCCCA GCTCACACCG CTTATGATCA ACTATATGCC GGCTTTTACC
                                                                            1020
      CAAGCTCACA CTTATACGCT GGCGGCCATC TACCCCTATG CTACTCAGCC TCAGGGAGAA
                                                                             1080
20
      TGGGCTTTCC AAGGTGAACT GCGTTACAAC TTTGCTCGCC GGACAGCTCT CGGTGGACGC
                                                                             1140
      TACGGTACCG GCTTGCGTAT CAACGTTTCG CATGTGCGTG GTCTGGACAA AAAGATGCTC
                                                                             1200
      AAAGAGAATC CCGACGAACT GATCGGAACG GATGGCTACA CCGTTTCTTT CTTCGGCATG
                                                                             1260
      GGCGACCTCT ATTATTCGGA TATAGATGTG GAGATTACTA AAAAGGTAAG CCCAGGATTC
                                                                             1320
      AACTTTACGC TCACCTACTT GAATCAGATC TACAATAACA AGGTACTGCA CGGTGCAGCC
                                                                             1380
25
      GGAGAGAGC CTGAGAAGAT CTATGCCAAT ATCTTCGTCT ATGATGGTAA GTATAAGCTG
                                                                             1440
      AGTAATAAGG TAGCCCTCCG TACCGAACTG CAATATTTGC ACACGAAGCA GGATCAGGGT
                                                                             1500
      GACTGGATCT ACGGCATGGC CGAGCTCTCT ATCCTGCTT CTCTGATGCT TTCCCTCTCG
                                                                             1560
      GAGCAGTATA ATATCGGAGA GACCAAGAAA CATTATGTCA TGGGGTCTGT CACCTATACT
                                                                             1620
      CACGGAGCAC ATCGAGTAGC TTTCTCTGCA GGCAAAACCC GTGCAGGGAT GAACTGCTCG
                                                                             1680
30
      GGAGGTGTAT GTCGTGTGGT CCCTGAGACT CAGGGATTCT ACCTTTCTTA TAGCACCAAT
                                                                             1740
                                                                             1743
      (2) INFORMATION FOR SEQ ID NO:244
35
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 717 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
40
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
45
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
50
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...7\overline{1}7
55
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:244
      ATGCGTTCTT TATTTTGAG CGCGTTGCGC AGCTCCTCTC TCCATGGTTC AGAGCGACGC
      AGTCGGATAA GTTCTTCTGT AGTCATGTCA ATAAGGCAGA AAATAAGGTT ATTCCATCTC
      TCGGTATGCG CCCAAACGCA TGATCATCTC ATCGAAATCC ACTTGGTGTG CATCGAATTC
60
      GGGGCCATCG ACACAGACGA ATTTCGTCTG TCCTCCCACG CTTATACGAC AAGCCCCACA
                                                                              240
      CATACCGGTG CCATCCACCA TAATTGTATT GAGAGAAGCT ATGGTCGGTA TCTCGTAACG
                                                                              300
      TTTGGTCAGG AGAGAAACGA ACTTCATCAT CACAGCCGGC CCGATCGTAA CGCAGAGGTC
                                                                             360
      TACCGTTTCC CGTTTGATAA CGCTTTCCAC TCCATCCGTT ACGAGGCCTT TCGTCCCATA
                                                                              420
      AGACCCATCG TCTGTCATGA TGATCACTTC ATCGCTATTG GCTCGCATTT GTTCTTCAAG
                                                                              480
65
      GATAACCAGA TCTTTAGTTC TGGCAGCCAA TACGACAATT ACACGGTTGC CTGCTTTGTG
                                                                             540
      GAAAGCCTCC ACGATCGGGA GCAAAGGAGC CACACCCACA CCGCCTCCGG CACAAACCAC
                                                                             600
      TGTGCCGACC TTTTCGATAT GCGTACTCTG TCCCAGCGGA CCTACCACAT CCGTGATATA
                                                                              660
      GTCGCCGACT TCGAGTTCGG CCAATTTCTT GGAAGATTTG CCCACGGCCT GAACCAC
70
      (2) INFORMATION FOR SEQ ID NO:245
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 633 base pairs
75
                 (B) TYPE: nucleic acid
```

```
(C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) NOLECULE TYPE: DNA (genomic)
 5
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
10
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROHOMAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
15
                 (B) LOCATION 1...633
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:245
      ATGTCAATAA GGCAGAAAAT AAGGTTATTC CATCTCTCGG TATGCGCCCA AACGCATGAT
                                                                                60
20
      CATCTCATCG AAATCCACTT GGTGTGCATC GAATTCGGGG CCATCGACAC AGACGAATTT
                                                                               120
      CGTCTGTCCT CCCACGCTTA TACGACAAGC CCCACACATA CCGGTGCCAT CCACCATAAT
                                                                               180
      TGTATTGAGA GAAGCTATGG TCGGTATCTC GTAACGTTTG GTCAGGAGAG AAACGAACTT
                                                                               240
      CATCATCACA GCCGGCCGA TCGTAACGCA GAGGTCTACC GTTTCCCGTT TGATAACGCT
                                                                               300
      TTCCACTCCA TCCGTTACGA GGCCTTTCGT CCCATAAGAC CCATCGTCTG TCATGATGAT
                                                                               360
25
      CACTTCATCG CTATTGGCTC GCATTTGTTC TTCAAGGATA ACCAGATCTT TAGTTCTGGC
                                                                               420
      AGCCAATACG ACAATTACAC GGTTGCCTGC TTTGTGGAAA GCCTCCACGA TCGGGAGCAA
                                                                               480
      AGGAGCCACA CCCACACCGC CTCCGGCACA AACCACTGTG CCGACCTTTT CGATATGCGT
                                                                               540
      ACTCTGTCCC AGCGGACCTA CCACATCCGT GATATAGTCG CCGACTTCGA GTTCGGCCAA
                                                                               600
      TTTCTTGGAA GATTTGCCCA CGGCCTGAAC CAC
                                                                               633
30
      (2) INFORMATION FOR SEQ ID NO:246
            (i) SEQUENCE CHARACTERISTICS:
35
                 (A) LENGTH: 2343 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
40
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) AHTI-SENSE: NO
45
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORTPHYROHOMAS GINGIVALIS
           (ix) FEATURE:
50
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...2343
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:246
      ATGGATCGTC CTAAGCCTTC ATATATTGTT CGAATAGCAG CCATTCTCTG CTTGTTTGTC
55
      GGCAGGCCTT TGTTTGCGCA GAGCTATGTG GACTACGTCG ATCCGCTGAT CGGGACGCTA AGTTCTTTTG AGCTGAGTGC GGGCAATACC TATCCGGTGA TCGGTTTACC GTGGGGAATG
                                                                               120
                                                                               180
      AATAGCTGGA CACCGATGAC CGGTGTACCC GGTGACGGCT GGCAATATAC CTACTCGGCA
                                                                               240
      CACAAGATTC GCGGATTCAA ACAGACCCAC CAACCCAGTC CTTGGATCAA CGACTACGGC
                                                                               300
60
      CAATTCTCCC TTCTTCCCCT TACGGCACCG CAGAAGCCAT CATCGAACGA CTCCATAGCT
                                                                               360
      CTGACTAAAT GGTGCAAGCA ACTCTTTTCG GACGAACAGA CCTCGTGGTT CTCGCACAAA
                                                                               420
      GCGGAGACGG CGACGCCATA CTATTATAGT GTCTATTTGG CCGATTACGA CACACGCGTG
                                                                               480
      GAGATGGCTC CGACCGAGCG TGCAGCTATC TTTCGCATAC GTTATTCCGG CAATACCGAA
                                                                               540
      AGTGGCTCCG GTCGATGGCT TCGTCTTGAT GCCTTTACCG GTGGTTCGGA GATTAGCATC
                                                                                600
65
      GTGGATCCTC ACACCGTAGT GGGCATATCT CGCAAGAATA GCGGAGGTGT GCCGGCTAAC
                                                                               660
      TTCGCCTGTT ATTTCATCCT GCAGTCCGAT ACTCCTATGG CCGATGTCCT GCTTGAGACA
                                                                               720
      GATACCGGCA AGTCAGACGA AGGCACAAGG GCATGGGCAG CCTGTCGCTT CGATTCGCAA
GAAGTTACCG TCCGGGTGGC ATCTTCTTT ATCAGTGTCG AGCAGGCCGA AAGAAATCTT
                                                                               780
                                                                               840
      GCGGAASTCA AAGGGCAGAS TTTCGACCGG ATCAGACTTG CCGGTCGCGA AGCTTGGAAT
                                                                               900
70
      AAGGTGCTCG GACGCATACA TGTGGAAGGA GGAACGAAGG ATGAGCGCAC TACATTCTAT
                                                                               960
      TCCGCACTCT ATCGCTGTCT GCTTTTTCCC CGTCGCTTCT ATGAGGAGGA TGCTTCCGGC
                                                                              1020
      AATTTTGTGC ATTACAGCCC CTACAATGGA GAGGTACTTC CCGGTTATCT CTATACCGAT
                                                                              1080
      ACCGGATTTT GGGACACTTT TCGAGCCCTT TTCCCCCTGC TCAATCTGCT GTATCCCGAT
                                                                              1140
      GAAAACATTA AAATTCAGGA AGGTCTGCTG AATGTATATC GCGAGAGTGG CTTTTTCCCC
                                                                              1200
75
      1260
```

```
GCGGATGCCT ACCTCAAGGG TGTTCGGGTA GAAGATACCC GTACACTGAT GAACGGACTC
                                                                               1320
      TTGCATGCTA CGAAAGCCGT CCATCCGAAA ATCTCCTCCA CGGGTCGCAA AGGTTGGGAG
                                                                               1380
      TGGTACAACT CCTTAGGTTA TGTTCCGGCT GATGCAGGCA TCGACGAAAG TGCTGCCCGT
                                                                               1440
      ACGCTCGAAT ATGCTTATAA CGATTGGTGC ATCCTCCGAC TGGGGCGCAC ATTGGGTTGG
                                                                               1500
 5
      GATAGAGCTG CATTGGACAC GTTGGCTCAT CGTTCGATGA ACTATCGTCA TCTGTTCGAT
                                                                               1560
      CCGGAAACCA AACTCATGCG CGGTAGAAAT CAGGATGGTA GTTTCCGGAC ACCTTTTTCC
                                                                               1620
      CCTTTCAAAT GGGGAGATGT ATTCACGGAG GGCAATGCCT GGCACTACAC TTGGTCGGTC
                                                                               1680
      TTTCATGATG TGCAGGGGCT TATCGACCTG ATGGGAGGAG ATCGCCCGTT CGTGTCTATG
                                                                               1740
      CTCGATTCGG TATTCAATAC TCCTCCTATG TTCGATGAGA GCTATTACGG ATTTGTCATC
                                                                               1800
10
      CACGAAATCA GAGAGATGCA AATAGCGGAT ATGGCCAATT ATGCTCATGG CAATCAACCC
                                                                               1860
      ATACAGCATA TGATATATCT GTATAATCAT GCCGGTCATC CATGGAAAGC TCAGGAGAGA
      CTACGCGAAG TGATGGGGCG GCTCTATCGT CCTACTCCGG ATGGGTATTG CGGCGATGAA GACAACGGAC AGACTTCGGC TTGGTACGTT TTCTCTGCTT TAGGCTTCTA TCCTGTTACA
                                                                               1980
                                                                               2040
      CCCGCTACGG ATCAGTATGT GCTCGGTTCG CCGATTTTTT CCAAGGTAAT ACTCTCTTTT
                                                                               2100
15
      CCCGACGGAC ACAAAACGGT GTTGCATGCT CCGGCCAACA GTGCCGATAC GCCTTACATC
                                                                               2160
      CGCTCGATCA GCGTAGAAGG AAAAGAATGG AGCTGCAATT ACCTGACTCA CGAACAGCTT
                                                                               2220
      CGCTCTTCTG CATCCATTCA ATGGATGATG GACACGAAAC CCAATTATAA TCGTGGTATG
                                                                               2280
      AAGGAAAGTG ACAGACCTTA TTCCTTCTCC ACGGAGCAAC AGCGTCGCGC TAATCACAGT
                                                                               2340
                                                                               2343
      TAA
20
      (2) INFORMATION FOR SEQ ID NO: 247
           (i) SEQUENCE CHARACTERISTICS:
25
                 (A) LENGTH: 813 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRAHDEDNESS: double
                 (D) TOPOLOGY: circular
30
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
35
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
40
                 (A) NAHE/KEY: misc_feature
                 (B) LOCATION 1...813
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:247
45
      ATGATGAAAT CCATGCGCAG CGTGCTGTTG CTACTCTTTC CATTGTCTTT GATCACTGCT
      TTGGGCTGTA GCAATAACAA AGCTGCCGAA TCGAAGTCTG TCTCTTTCGA TTCGGCCTAT
                                                                                120
      CTCGAACGCT ACATCCCTCT GCGGGCAGAC ATAGATACGC CATCGCTGCA TGTGATGATC
                                                                                180
      AGCTACGTCT ATCCTTCGGG AGATGATATG CTCACAGAGA TTTTCAACGG TTTGCTCTTC
                                                                                240
      GGCGACAGCC TGATGGATTC CTCTTCGCCG GAGAATGCCA TGGAAGGCTA TGCACAGATG
                                                                                300
50
      CTGGGAGAAG ACTATCGCTC TAACAATGCC GAAGCCAATC TGCAAGGGCT TCCTTCTGAC
                                                                                360
      CTTTTGGACT ATATCTACAA GCAGGAAAAT ACCATCGCTT ATTGCGATAC GGGATTGATC
                                                                                420
      TCCACGCGCA TCAATACATA TACTTACGAA GGCGGTGCAC ATACGGAGAA TACAGTCCGG
                                                                                480
      TTTGCCAACA TCCTTCGCAC CACCGGCAAG GTGCTCGAAG AGCGAGATAT ATTCAAGATC
                                                                                540
      GACTATGCGG AAAGGCTGTC CGCACTCATC ATAGGACAAT TGGTGCACGA TTTCGGCAAG
                                                                                600
55
      ACCACACCTG CCGAATTGGA TGCAATAGGT TTCTTCAACG CAGAAGAAAT ACAGCCCAAT
                                                                                660
      GGCAATTTTA TGATCGATGA CAAAGGTCTC ACATACTGTT TCAATGAGTA TCAGATAGCT
                                                                                720
      GCTTATGCCA GAGGTGCTGT CTATGTCCGT CTCGGATATG ACGTATTGGC TCCTTTGCTA
                                                                                780
      AGGGATGATT CCCCACTAAA GCGTTACTTG CCG
60
      (2) INFORMATION FOR SEQ ID NO: 248
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 810 base pairs
65
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
70
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
75
           (vi) ORIGINAL SOURCE:
```

```
(A) ORGANISM: PORYPHYROMONAS GINGIVALIS
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
 5
                  (B) LOCATION 1...8\overline{1}0
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:248
       ATGAAATCCA TGCGCAGCGT GCTGTTGCTA CTCTTTCCAT TGTCTTTGAT CACTGCTTTG
                                                                                    60
10
       GGCTGTAGCA ATAACAAAGC TGCCGAATCG AAGTCTGTCT CTTTCGATTC GGCCTATCTC
                                                                                   120
       GAACGCTACA TCCCTCTGCG GGCAGACATA GATACGCCAT CGCTGCATGT GATGATCAGC
                                                                                   180
       TACGTCTATC CTTCGGGAGA TGATATGCTC ACAGAGATTT TCAACGGTTT GCTCTTCGGC
                                                                                   240
       GACAGCCTGA TGGATTCCTC TTCGCCGGAG AATGCCATGG AAGGCTATGC ACAGATGCTG
                                                                                   300
       GGAGAAGACT ATCGCTCTAA CAATGCCGAA GCCAATCTGC AAGGGCTTCC TTCTGACCTT
                                                                                   360
15
       TTGGACTATA TCTACAAGCA GGAAAATACC ATCGCTTATT GCGATACGGG ATTGATCTCC
                                                                                   420
       ACGCGCATCA ATACATATAC TTACGAAGGC GGTGCACATA CGGAGAATAC AGTCCGGTTT
GCCAACATCC TTCGCACCAC CGGCAAGGTG CTCGAAGAGC GAGATATATT CAAGATCGAC
                                                                                   480
                                                                                   540
       TATGCGGAAA GGCTGTCCGC ACTCATCATA GGACAATTGG TGCACGATTT CGGCAAGACC ACACCTGCCG AATTGGATGC AATAGGTTTC TTCAACGCAG AAGAAATACA GCCCAATGGC
                                                                                   600
                                                                                   660
20
       AATTTTATGA TCGATGACAA AGGTCTCACA TACTGTTTCA ATGAGTATCA GATAGCTGCT
                                                                                   720
       TATGCCAGAG GTGCTGTCTA TGTCCGTCTC GGATATGACG TATTGGCTCC TTTGCTAAGG
                                                                                   780
       GATGATTCCC CACTAAAGCG TTACTTGCCG
25
       (2) INFORMATION FOR SEQ ID NO: 249
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 801 base pairs
                  (B) TYPE: nucleic acid
30
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
35
          (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
40
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...8\overline{0}1
45
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:249
      ATGCGCAGCG TGCTGTTGCT ACTCTTTCCA TTGTCTTTGA TCACTGCTTT GGGCTGTAGC
      AATAACAAAG CTGCCGAATC GAAGTCTGTC TCTTTCGATT CGGCCTATCT CGAACGCTAC
                                                                                   120
50
      ATCCCTCTGC GGGCAGACAT AGATACGCCA TCGCTGCATG TGATGATCAG CTACGTCTAT
      CCTTCGGGAG ATGATATGCT CACAGAGATT TTCAACGGTT TGCTCTTCGG CGACAGCCTG
                                                                                   240
      ATGGATTCCT CTTCGCCGGA GAATGCCATG GAAGGCTATG CACAGATGCT GGGAGAAGAC
                                                                                   300
      TATCGCTCTA ACAATGCCGA AGCCAATCTG CAAGGGCTTC CTTCTGACCT TTTGGACTAT
                                                                                   360
      ATCTACAAGC AGGAAAATAC CATCGCTTAT TGCGATACGG GATTGATCTC CACGCGCATC
                                                                                   420
55
      AATACATATA CTTACGAAGG CGGTGCACAT ACGGAGAATA CAGTCCGGTT TGCCAACATC
                                                                                   480
       CTTCGCACCA CCGGCAAGGT GCTCGAAGAG CGAGATATAT TCAAGATCGA CTATGCGGAA
                                                                                  540
      AGGCTGTCCG CACTCATCAT AGGACAATTG GTGCACGATT TCGGCAAGAC CACACCTGCC
                                                                                  600
      GAATTGGATG CAATAGGTTT CTTCAACGCA GAAGAAATAC AGCCCAATGG CAATTTTATG
                                                                                  660
      ATCGATGACA AAGGTCTCAC ATACTGTTTC AATGAGTATC AGATAGCTGC TTATGCCAGA
                                                                                  720
60
      GGTGCTGTCT ATGTCCGTCT CGGATATGAC GTATTGGCTC CTTTGCTAAG GGATGATTCC
                                                                                  780
      CCACTAAAGC GTTACTTGCC G
                                                                                  801
       (2) INFORMATION FOR SEQ ID NO:250
65
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 777 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
70
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
75
```

```
(iv) ANTI-SEUSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
 5
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...777
10
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:250
       ATGAAAAAA CTACTTTGAC AGGATCGATA TGTGCTTTAC TCCTGTTTTT GGGTCTCTCG
       GCCAATGCCC AATCGAAGTT AAAGATCAAG AGCATTGAGG CAGCTACCAC TTTCAGTTCG
                                                                                    120
       GCCACGGCCG GAAATGGTTT TGGTGGCAAT ATCTTCGGCA TGGACATGAG CATACGGATG
                                                                                    180
15
       AGGGTACACC ACAGCATTCT GCCCGAAGGG TTGGATTTTT CGGTAGGAAT ACATGAAAGA
                                                                                    240
       AGAGCACACT GGGAAGAGGC CGGAAGTCCG AAGCTCATGT ATACGAATGT CCCAAGTATC
                                                                                    300
       ATTGGTATTG TTGAAAAGGT AATAGTCTTC GAAGACGCAG AAGACTTTTT TGACAAAAAA
                                                                                    360
       GCTCTCGGCC GCTTCCTCAT CAGTTTGGGG ATATCCTATA CCAAGCATCT GGGAGCGTAT
                                                                                    420
       TGGGGATGGA CCAATGACGC CCATATTCTT TTCTCACCGA TACCCAAGAG CAAGGTCCAC TATGACACCT ACACAAGAGC TGGCAGTGAC CTTGTACTTC AGTCCGAAGA TGTTGCCACA
                                                                                    480
20
                                                                                    540
       GTGAGCAATG GCTTTTCACC GGGGATCGGA CTCAAAAGTT CTATTTGGTG GAAAATGCCC
                                                                                    600
       ATCAGGGGA AATATGATTT TCGCCTCGGT TTCAGCCTGG GCTATGAGTA TCTGAACCTG
                                                                                    660
       CTATATCCGT ATCGTAATTT CAAGCTGGAT GGAAATAAGC CGCTTTCAGC ACTATCTCCT
                                                                                    720
       CGCATGAACC ACATCGGCCA TGTGGGCTTC AACTTTACCG TGGGTCTTTG GACTAAT
                                                                                    777
25
       (2) INFORMATION FOR SEQ ID NO:251
            (i) SEQUENCE CHARACTERISTICS:
30
                  (A) LENGTH: 3798 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
35
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
           (iv) ANTT-SENSE: NO
40
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
45
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...3798
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:251
50
       ATGGGCAAGT ATAAAAGAGC TAAGTACCGC TATTGGCTTT TTCCTTTCTG TTCGGATTAT
       TATACCTTTG AGGGAGTTAC TTTTTTATGC GCATCTGACG ATATGACAAC CAAGAAACCC
                                                                                    120
       CAAGCCATTT TAGACTTAGA GAAGGCCTAT AACATTGAAA TTCCTGATCT CTCCTCACAA
       GAAGGGATAA GCTGGTCGGT AAATAGATAT TTCAAGCAAG ATTCCTCCGG TGCAGTCGTT
                                                                                    240
       GAGCTTTGCT TGCGAGAATG TCAGATAGAA AGCATGACTT GGCTTATTGA TTTTCCTGCT
                                                                                    300
55
       CTAAAAAAGC TTGATCTATC GTATAACCAA ATCAGTAAGC TAGAGGGTCT AGAACGTCTT
                                                                                    360
       ACTTCGTTAA CAAAACTTCG TCTAAGAAGT AACCAAATCC GTAAACTAGA GGGCCTGGAT
                                                                                    420
       AGTCTCACCT CGCTAACAAA ACTTTCTCTC TCCGATAACC AAATCAGTAA GCTAGAGGGT
CTGGAACGTC TCACCTCGTT AGCGGAGCTT TATCTTTTGG ATAACCAAAT CAGTAAACTA
                                                                                    480
                                                                                    540
       GAGGGTCTGG AACGTCTCAC GTCCTTAGCA ACGCTTGAAC TATCGGGTAA CCAAATCCGT
AAGCTGGAGG GTCTGGAACG TCTCACGTCC TTAGCAACGC TTGAACTATC GGGTAACCAA
                                                                                    600
60
                                                                                    660
       ATCCGTAAGC TAGAGGGTCT GGAACGTCTC ACTTCGTTAA CAAAGCTTCG TCTAAGAAGT
                                                                                    720
       AACCAAATCA GTAAGCTAGA GGGTCTGGAA CGTCTCACGT CCTTAGCAAC GCTTGAACTA
                                                                                    780
       TCGGGTAACC AAATCCGTAA GCTGGAGGGT CTGGAACGTC TCACGTCCTT AGCAACGCTT
                                                                                    840
       GAACTGTCGG GTAACCAAAT CAGTAAGCTA GAGGGTCTGG AACGTCTCTC TTCGTTAACA
                                                                                    900
65
       AAGCTTCGTC TAAGAAGTAA CCAGATCAGT AAACTAGAGG GCCTGGAACG TCTCACCTCG
       CTAACAAAAC TTTCTCTCTC CGATAACCAA ATCAGTAAGC TAGAGGGTCT GGAACGTCTC
                                                                                   1020
       ACCTCGTTAG CGGAGCTTTA TCTTTTGGAT AACCAAATCC GTAAGCTGGA GGGCCTGGAA
       CGTCTCACCT CGTTAACAAA GCTTCGTCTA AGAAGTAACC AAATCAGTAA ACTAGAGGGC
                                                                                   1140
       CTGGATAGTC TCACCTCGCT AACAAAACTT TCTCTCTCCG ATAACCAAAT CAGTAAACTA
                                                                                   1200
70
       GAGGGCCTGG AACGTCTCAC GTCCTTAGCG GAGCTTTATC TTTTGGATAA CCAAATCCGT
AAGCTGGAGG GTCTTGATGG TCTTGCTTCC TTAACAAGGC TTAGTCTAAG GCGCAACCAA
                                                                                   1260
                                                                                   1320
       ATCAGTAAGC TGGAAGGACT AGACAGACTA AAGGTTTTGA GAAAACTTGA TGTTTCGGGC
                                                                                   1380
       AATGATATTC AATCTATTGA TGATATTAAG CTATTGGCTC CGATTCTGGA GCAAACTTTA
                                                                                   1440
       GAAAAACTGA GAATCCATGA CAATCCATTT GTTGCATCAT CAGGCTTGAT ACTCTCTCCT
                                                                                   1500
75
       TATGATAATC ATTTGCCGGA GATTAAAGCT CTTCTTGAAA AAGAAAAAGA AAAACAGAAA
                                                                                   1560
```

```
AAGACTTCAG TTGAATATCA CCCATTTTGC AAAGTAATGC TATTGGGAAA TCATTCTTCG
                                                                             1620
      GGTAAAACAA CATTTCTTAG TCAATACGAT ACAAATTATA CGTATCAGAA AAATACACAT
                                                                             1680
      GTGTTGTCGA TACATCGAAG CAATAACCCT AATGCGATCT TTTACGACTT TGGGGGACAG
                                                                             1740
      GACTATTATC ATGGGATTTA CCAAGCCTTT TTTACCACCC AATCGTTATA CCTTCTTTT
                                                                             1800
 5
      TGGGATGCTA AGAAGGATCG AAACTTTGTG AGCGTAGATG ATAAAGAATA TCAGACTCTT
                                                                             1860
      AATTTCAATC GCCCCTATTG GTTAGGACAG ATAGCCTATG CCTGCAATCG TTGTATGTCC
                                                                             1920
      GTTGGAGGAA ATCCTGATGG CAAGGACACA CCACAGACCA CAGACGATAC AATTATCATT
                                                                             1980
      CAGACTCATG CCGATGAAAC GGGCGCTAAG CAGCAAACCT TAGGCTGTGC AGCCGAGAAT
      GGAGTATTGG AAGAAATCTA TGTATCCTTA GAGCCCAAGG CGAATAGTGC CGTACATGCG
                                                                             2100
10
      CTCAACTATC TGAATGAGCG GGTGCGAGAA GTTGTCGCAA GCAGGAGTAA ATCAATTCAG
                                                                             2160
      ATCACAGAAA AAGATAAGGG ATTGTACGAA GCTCTTCCCA CAATCGCCGG TGATAATAAA
                                                                             2220
      CACATCCCTA TCTCTCGA AGCTCTTGCG GCTCAATTGA ATAAGGGAAG AGCTGAAAAT
                                                                             2280
      GATCTTTACA CCATAGAGTA TCTACAGACC GAATTGAACC AGCTTAGTCT GCGAGGGGAG
                                                                             2340
      GTGCTTTACT ATCGTGAGAA TGAGAAGCTG AACAATTATG TCTGGTTAGA TCCGGCAGCT
                                                                             2400
      TTTGTCCAAA TGATTCATGG AGAAATCCTC CAAAAAGACA ACATCAATAG AGGAACAGTT
15
                                                                             2460
      CCTAAAGACA TTTTTGAATG CAAACTGCAT AATCTAAGTT CCGGAAGTAT ATTTGAAGAA
                                                                             2520
      GATGGCCAAA ATGGTAATAT GATCTTGCAG CTATTATTGG AAGAGCTGAT CGTATATGAA
                                                                             2580
      GATAAGSACT GCTATGTGAT ACCGGGCTAT CTCCCTTTGC ATTCCGATGA CGAAGCCTAT
                                                                             2640
      AAATGGCTTA CTTTGGGATT CGAGAGGCCC AATTTTGTCC TCAAATTCGA ACGTTTTATC
                                                                             2700
20
      CCCTTTGGCC TGATCAACCA GATTATAGCC TACTATGGCC GGGAAGAAGG TGCTCTAAAG
                                                                             2760
      CGGTATTGGC GAGATCAGGT CATCTTCACA GCAGGCCGTG AGATGGATAG GCAAACGCTT
                                                                             2820
      GAGCAAGAAG AAGAGAAAGA GGGTTTGCCC AAGACGAATG CCGAGGATTA TCAGATCTGG
      ATCAAGCTCG ACTTTACCGA CTTGGCCATA TCCGTATTCA TCAAAGAGCA GAGAAAGACA
                                                                             2940
      TCAGCTAAGG ATATGCAGCG GAAAGAGGCT ACTATCCTCA GTGATATGTT GGATATGTAT
                                                                             3000
25
      TGGAACAATA TCCCTCCGAG GGAGCAAATA GGAGATAAGG ATACGGAGCA AACGAGAAGC
                                                                             3060
      ACTATTCGTG AAACAACAG AAAGAAGAG CCCATCCAGG ATCTCTACCT CTCCTGTGCC
                                                                             3120
      CAAGCGGATA AAGATTTGAC GGAGTCTCAT TATATCCATT TGGGCACGCT GGACGATGAA AGCAAGACTA CGGCGAGGAT TGCAGCCTAT CCGTTGAAGA ACGGCGTTAT CGATAAAGAG
                                                                             3180
                                                                             3240
      CGGGTGCGAG AAGTATCGAC TCGTCCCTAC AAACATCTTT CCGTCAATAA AAATCTGGCT
                                                                             3300
30
      ACTGCAAAAC AGATCTTTAT TTCCTATTCC AAAGAGGATC AGACTGAACT GGAGACCTGT
                                                                             3360
      CTGCAATTTT TCAAACCCTT GGAGAAGAAT GGTCAGATCG AGATCTACTA TGATAAGTTG
                                                                             3420
      ACTAAGTTTG AAACACCTAT TCACCCTGAA ATAAGAAAGC GTATTGTCGA AGCCGACTGT
                                                                             3480
      ATAATCGCTT TGATCAGCCA ACGCTATCTG GCCACGGATT ACATCCTGGA TCATGAGTTG
                                                                             3540
      CCTGTATTTC GGGAGTATAA CAAGACCATA GTGCCGATAT TGATCAAGCC TTGTACATTC
                                                                             3600
35
      GAAGACGATG AGTTCCTTCG GGAGAAATAT TTTGCTCAGA AAGCTCAAAT AATCAATCTT
                                                                             3660
      GGAAAAGAGG GAAAAACCAT TAAAGCTTAT GATAGTATTA CGGCATCAGC CCATCGTGAT
                                                                             3720
      GAAAATTGGG TGGCAGTAGT CAGAGAGTTC AAAGAGAAA TATTAAGAAT AACAAAACAG
                                                                             3780
      GAGGTAAATA CAGATGAA
                                                                              3798
40
      (2) INFORMATION FOR SEQ ID NO:252
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 3696 base pairs
45
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
50
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SENSE: NO
55
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
60
                 (B) LOCATION 1...3696
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:252
      ATGACAACCA AGAAACCCCA AGCCATTTTA GACTTAGAGA AGGCCTATAA CATTGAAATT
                                                                               60
65
      CCTGATCTCT CCTCACAGA AGGGATAAGC TGGTCGGTAA ATAGATATTT CAAGCAAGAT
                                                                              120
      TCCTCCGGTG CAGTCGTTGA GCTTTGCTTG CGAGAATGTC AGATAGAAAG CATGACTTGG
                                                                              180
      CTTATTGATT TTCCTGCTCT AAAAAAGCTT GATCTATCGT ATAACCAAAT CAGTAAGCTA
                                                                              240
      GAGGGTCTAG AACGTCTTAC TTCGTTAACA AAACTTCGTC TAAGAAGTAA CCAAATCCGT
                                                                              300
      AAACTAGAGG GCCTGGATAG TCTCACCTCG CTAACAAAAC TTTCTCTCTC CGATAACCAA
                                                                              360
70
      ATCAGTAAGC TAGAGGGTCT GGAACGTCTC ACCTCGTTAG CGGAGCTTTA TCTTTTGGAT
                                                                              420
      AACCAAATCA GTAAACTAGA GGGTCTGGAA CGTCTCACGT CCTTAGCAAC GCTTGAACTA
                                                                              480
      TCGGGTAACC AAATCCGTAA GCTGGAGGGT CTGGAACGTC TCACGTCCTT AGCAACGCTT
                                                                              540
      GAACTATCGG GTAACCAAAT CCGTAAGCTA GAGGGTCTGG AACGTCTCAC TTCGTTAACA
                                                                              600
      AAGCTTCGTC TAAGAAGTAA CCAAATCAGT AAGCTAGAGG GTCTGGAACG TCTCACGTCC
                                                                              660
75
      TTAGCAACGC TTGAACTATC GGGTAACCAA ATCCGTAAGC TGGAGGGTCT GGAACGTCTC
                                                                              720
```

```
ACGTCCTTAG CAACGCTTGA ACTGTCGGGT AACCAAATCA GTAAGCTAGA GGGTCTGGAA
                                                                                      780
       CGTCTCTCTT CGTTAACAAA GCTTCGTCTA AGAAGTAACC AGATCAGTAA ACTAGAGGGC
                                                                                      840
       CTGGAACGTC TCACCTCGCT AACAAAACTT TCTCTCTCCG ATAACCAAAT CAGTAAGCTA
                                                                                      900
       GAGGGTCTGG AACGTCTCAC CTCGTTAGCG GAGCTTTATC TTTTGGATAA CCAAATCCGT
                                                                                      960
  5
       AAGCTGGAGG GCCTGGAACG TCTCACCTCG TTAACAAAGC TTCGTCTAAG AAGTAACCAA
                                                                                    1020
       ATCAGTAAAC TAGAGGGCCT GGATAGTCTC ACCTCGCTAA CAAAACTTTC TCTCTCCGAT
                                                                                    1080
       AACCAAATCA GTAAACTAGA GGGCCTGGAA CGTCTCACGT CCTTAGCGGA GCTTTATCTT
                                                                                    1140
       TTGGATAACC AAATCCGTAA GCTGGAGGGT CTTGATGGTC TTGCTTCCTT AACAAGGCTT
                                                                                    1200
       AGTCTAAGGC GCAACCAAAT CAGTAAGCTG GAAGGACTAG ACAGACTAAA GGTTTTGAGA
                                                                                    1260
10
       AAACTTGATG TTTCGGGCAA TGATATTCAA TCTATTGATG ATATTAAGCT ATTGGCTCCG
                                                                                    1320
       ATTCTGGAGC AAACTTAGA AAAACTGAGA ATCCATGACA ATCCATTTGT TGCATCATCA
GGCTTGATAC TCTCTCCTTA TGATAATCAT TTGCCGGAGA TTAAAGCTCT TCTTGAAAAA
                                                                                    1380
                                                                                    1440
       GAAAAAGAAA AACAGAAAAA GACTTCAGTT GAATATCACC CATTTTGCAA AGTAATGCTA
                                                                                    1500
       TTGGGAAATC ATTCTTCGGG TAAAACAACA TTTCTTAGTC AATACGATAC AAATTATACG
                                                                                    1560
15
       TATCAGAAAA ATACACATGT GTTGTCGATA CATCGAAGCA ATAACCCTAA TGCGATCTTT
                                                                                    1620
       TACGACTTTG GGGGACAGGA CTATTATCAT GGGATTTACC AAGCCTTTTT TACCACCCAA
                                                                                    1680
       TCGTTATACC TTCTCTTTTG GGATGCTAAG AAGGATCGAA ACTTTGTGAG CGTAGATGAT
                                                                                    1740
       AAAGAATATC AGACTCTTAA TTTCAATCGC CCCTATTGGT TAGGACAGAT AGCCTATGCC
                                                                                    1800
       TGCAATCGTT GTATGTCCGT TGGAGGAAAT CCTGATGGCA AGGACACAC ACAGACCACA
                                                                                    1860
20
       GACGATACAA TTATCATTCA GACTCATGCC GATGAAACGG GCGCTAAGCA GCAAACCTTA
                                                                                    1920
       GGCTGTGCAG CCGAGAATGG AGTATTGGAA GAAATCTATG TATCCTTAGA GCCCAAGGCG
                                                                                    1980
       AATAGTGCCG TACATGCGCT CAACTATCTG AATGAGCGGG TGCGAGAAGT TGTCGCAAGC AGGAGTAAAT CAATTCAGAT CACAGAAAAA GATAAGGGAT TGTACGAAGC TCTTCCCACA
                                                                                    2040
                                                                                    2100
       ATCGCCGGTG ATAATAAACA CATCCCTATC TCTCTCGAAG CTCTTGCGGC TCAATTGAAT
                                                                                    2160
25
       AAGGGAAGAG CTGAAAATGA TCTTTACACC ATAGAGTATC TACAGACCGA ATTGAACCAG
                                                                                    2220
       CTTAGTCTGC GAGGGGAGGT GCTTTACTAT CGTGAGAATG AGAAGCTGAA CAATTATGTC
                                                                                    2280
       TGGTTAGATC CGGCAGCTTT TGTCCAAATG ATTCATGGAG AAATCCTCCA AAAAGACAAC
                                                                                    2340
       ATCAATAGAG GAACAGTTCC TAAAGACATT TTTGAATGCA AACTGCATAA TCTAAGTTCC
                                                                                    2400
       GGAAGTATAT TTGAAGAAGA TGGCCAAAAT GGTAATATGA TCTTGCAGCT ATTATTGGAA
                                                                                    2460
30
       GAGCTGATCG TATATGAAGA TAAGGACTGC TATGTGATAC CGGGCTATCT CCCTTTGCAT
                                                                                    2520
       TCCGATGACG AAGCCTATAA ATGGCTTACT TTGGGATTCG AGAGGCCCAA TTTTGTCCTC
                                                                                    2580
       AAATTCGAAC GTTTTATCCC CTTTGGCCTG ATCAACCAGA TTATAGCCTA CTATGGCCGG
                                                                                    2640
       GAAGAAGGTG CTCTAAAGCG GTATTGGCGA GATCAGGTCA TCTTCACAGC AGGCCGTGAG
                                                                                    2700
       ATGGATAGGC AAACGCTTGA GCAAGAAGAA GAGAAAGAGG GTTTGCCCAA GACGAATGCC
                                                                                    2760
35
       GAGGATTATC AGATCTGGAT CAAGCTCGAC TTTACCGACT TGGCCATATC CGTATTCATC
                                                                                    2820
       AAAGAGCAGA GAAAGACATC AGCTAAGGAT ATGCAGCGGA AAGAGGCTAC TATCCTCAGT
                                                                                    2880
       GATATGTTGG ATATGTATTG GAACAATATC CCTCCGAGGG AGCAAATAGG AGATAAGGAT
ACGGAGCAAA CGAGAAGCAC TATTCGTGAA ACAAACAGAA AGAAGAGACC CATCCAGGAT
                                                                                    2940
                                                                                    3000
       CTCTACCTCT CCTGTGCCCA AGCGGATAAA GATTTGACGG AGTCTCATTA TATCCATTTG
                                                                                    3060
40
       GGCACGCTGG ACGATGAAAG CAAGACTACG GCGAGGATTG CAGCCTATCC GTTGAAGAAC
                                                                                    3120
       GGCGTTATCG ATAAAGAGCG GGTGCGAGAA GTATCGACTC GTCCCTACAA ACATCTTTCC
GTCAATAAAA ATCTGGCTAC TGCAAAACAG ATCTTTATTT CCTATTCCAA AGAGGATCAG
                                                                                    3180
                                                                                    3240
       ACTGAACTGG AGACCTGTCT GCAATTTTTC AAACCCTTGG AGAAGAATGG TCAGATCGAG
ATCTACTATG ATAAGTTGAC TAAGTTTGAA ACACCTATTC ACCCTGAAAT AAGAAAGCGT
                                                                                    3300
                                                                                    3360
45
       ATTGTCGAAG CCGACTGTAT AATCGCTTTG ATCAGCCAAC GCTATCTGGC CACGGATTAC
                                                                                    3420
       ATCCTGGATC ATGAGTTGCC TGTATTTCGG GAGTATAACA AGACCATAGT GCCGATATTG
                                                                                    3480
       ATCAAGCCTT GTACATTCGA AGACGATGAG TTCCTTCGGG AGAAATATTT TGCTCAGAAA
                                                                                    3540
       GCTCAAATAA TCAATCTTGG AAAAGAGGGA AAAACCATTA AAGCTTATGA TAGTATTACG
                                                                                    3600
       GCATCAGCCC ATCGTGATGA AAATTGGGTG GCAGTAGTCA GAGAGTTCAA AGAGAAGATA
                                                                                    3660
50
       TTAAGAATAA CAAAACAGGA GGTAAATACA GATGAA
                                                                                    3696
       (2) INFORMATION FOR SEQ ID NO: 253
55
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 3525 base pairs
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: circular
60
            (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
65
            (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: PORYPHYROHONAS GINGIVALIS
70
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...3525
           (XI) SEQUENCE DESCRIPTION: SEQ ID NO:253
75
```

```
ATGACTTGGC TTATTGATTT TCCTGCTCTA AAAAAGCTTG ATCTATCGTA TAACCAAATC
      AGTAAGCTAG AGGGTCTAGA ACGTCTTACT TCGTTAACAA AACTTCGTCT AAGAAGTAAC
      CAAATCCGTA AACTAGAGGG CCTGGATAGT CTCACCTCGC TAACAAAACT TTCTCTCTCC
      GATAACCAAA TCAGTAAGCT AGAGGGTCTG GAACGTCTCA CCTCGTTAGC GGAGCTTTAT
                                                                             240
      CTTTTGGATA ACCAAATCAG TAAACTAGAG GGTCTGGAAC GTCTCACGTC CTTAGCAACG
                                                                             300
      CTTGAACTAT CGGGTAACCA AATCCGTAAG CTGGAGGGTC TGGAACGTCT CACGTCCTTA
                                                                             360
      GCAACGCTTG AACTATCGGG TAACCAAATC CGTAAGCTAG AGGGTCTGGA ACGTCTCACT
                                                                             420
      TCGTTAACAA AGCTTCGTCT AAGAAGTAAC CAAATCAGTA AGCTAGAGGG TCTGGAACGT
                                                                             480
      CTCACGTCCT TAGCAACGCT TGAACTATCG GGTAACCAAA TCCGTAAGCT GGAGGGTCTG
                                                                             540
10 .
     GAACGTCTCA CGTCCTTAGC AACGCTTGAA CTGTCGGGTA ACCAAATCAG TAAGCTAGAG
                                                                             600
      GGTCTGGAAC GTCTCTCTC GTTAACAAAG CTTCGTCTAA GAAGTAACCA GATCAGTAAA
                                                                             660
      CTAGAGGGCC TGGAACGTCT CACCTCGCTA ACAAAACTTT CTCTCCCGA TAACCAAATC
                                                                             720
      AGTAAGCTAG AGGGTCTGGA ACGTCTCACC TCGTTAGCGG AGCTTTATCT TTTGGATAAC
                                                                             780
      CAAATCCGTA AGCTGGAGGG CCTGGAACGT CTCACCTCGT TAACAAAGCT TCGTCTAAGA
                                                                             810
15
      AGTAACCAAA TCAGTAAACT AGAGGGCCTG GATAGTCTCA CCTCGCTAAC AAAACTTTCT
                                                                             900
      CTCTCCGATA ACCAAATCAG TAAACTAGAG GGCCTGGAAC GTCTCACGTC CTTAGCGGAG
      CTTTATCTTT TGGATAACCA AATCCGTAAG CTGGAGGGTC TTGATGGTCT TGCTTCCTTA
                                                                            1020
      ACAAGGCTTA GTCTAAGGCG CAACCAAATC AGTAAGCTGG AAGGACTAGA CAGACTAAAG
                                                                            1080
      GTTTTGAGAA AACTTGATGT TTCGGGCAAT GATATTCAAT CTATTGATGA TATTAAGCTA
                                                                            1140
20
      TTGGCTCCGA TTCTGGAGCA AACTTTAGAA AAACTGAGAA TCCATGACAA TCCATTTGTT
      GCATCATCAG GCTTGATACT CTCTCCTTAT GATAATCATT TGCCGGAGAT TAAAGCTCTT
                                                                            1260
      CTTGAAAAAG AAAAAGAAAA ACAGAAAAAG ACTTCAGTTG AATATCACCC ATTTTGCAAA
                                                                            1320
      GTAATGCTAT TGGGAAATCA TTCTTCGGGT AAAACAACAT TTCTTAGTCA ATACGATACA
                                                                            1380
      AATTATACGT ATCAGAAAAA TACACATGTG TTGTCGATAC ATCGAAGCAA TAACCCTAAT
                                                                            1440
25
      GCGATCTTTT ACGACTTTGG GGGACAGGAC TATTATCATG GGATTTACCA AGCCTTTTTT
                                                                            1500
      ACCACCCAAT CGTTATACCT TCTCTTTTGG GATGCTAAGA AGGATCGAAA CTTTGTGAGC
                                                                            1560
      GTAGATGATA AAGAATATCA GACTCTTAAT TTCAATCGCC CCTATTGGTT AGGACAGATA
                                                                            1620
      GCCTATGCCT GCAATCGTTG TATGTCCGTT GGAGGAAATC CTGATGGCAA GGACACCA
                                                                            1680
      CAGACCACAG ACGATACAAT TATCATTCAG ACTCATGCCG ATGAAACGGG CGCTAAGCAG
                                                                            1740
30
      CAAACCTTAG GCTGTGCAGC CGAGAATGGA GTATTGGAAG AAATCTATGT ATCCTTAGAG
                                                                            1800
      CCCAAGGCGA ATAGTGCCGT ACATGCGCTC AACTATCTGA ATGAGCGGGT GCGAGAAGTT
                                                                            1860
      GTCGCAAGCA GGAGTAAATC AATTCAGATC ACAGAAAAAG ATAAGGGATT GTACGAAGCT
                                                                            1920
      CTTCCCACAA TCGCCGGTGA TAATAAACAC ATCCCTATCT CTCTCGAAGC TCTTGCGGCT
                                                                            1980
      CAATTGAATA AGGGAAGAGC TGAAAATGAT CTTTACACCA TAGAGTATCT ACAGACCGAA
                                                                            2040
35
      TTGAACCAGC TTAGTCTGCG AGGGGAGGTG CTTTACTATC GTGAGAATGA GAAGCTGAAC
                                                                            2100
      AATTATGTCT GGTTAGATCC GGCAGCTTTT GTCCAAATGA TTCATGGAGA AATCCTCCAA
                                                                            2160
      AAAGACAACA TCAATAGAGG AACAGTTCCT AAAGACATTT TTGAATGCAA ACTGCATAAT
      CTAAGTTCCG GAAGTATATT TGAAGAAGAT GGCCAAAATG GTAATATGAT CTTGCAGCTA
                                                                            2280
      TTATTGGAAG AGCTGATCGT ATATGAAGAT AAGGACTGCT ATGTGATACC GGGCTATCTC
40
      CCTTTGCATT CCGATGACGA AGCCTATAAA TGGCTTACTT TGGGATTCGA GAGGCCCAAT
                                                                            2400
      TTTGTCCTCA AATTCGAACG TTTTATCCCC TTTGGCCTGA TCAACCAGAT TATAGCCTAC
                                                                            2460
      TATGGCCGGG AAGAAGGTGC TCTAAAGCGG TATTGGCGAG ATCAGGTCAT CTTCACAGCA
      GGCCGTGAGA TGGATAGGCA AACGCTTGAG CAAGAAGAAG AGAAAGAGGG TTTGCCCAAG
                                                                            2580
      ACGAATGCCG AGGATTATCA GATCTGGATC AAGCTCGACT TTACCGACTT GGCCATATCC
                                                                            2640
45
      GTATTCATCA AAGAGCAGAG AAAGACATCA GCTAAGGATA TGCAGCGGAA AGAGGCTACT
                                                                            2700
      ATCCTCAGTG ATATGTTGGA TATGTATTGG AACAATATCC CTCCGAGGGA GCAAATAGGA
                                                                            2760
      GATAAGGATA CGGAGCAAAC GAGAAGCACT ATTCGTGAAA CAAACAGAAA GAAGAGACCC
                                                                            2820
      ATCCAGGATC TCTACCTCTC CTGTGCCCAA GCGGATAAAG ATTTGACGGA GTCTCATTAT
                                                                            2880
      ATCCATTIGG GCACGCIGGA CGATGAAAGC AAGACTACGG CGAGGATIGC AGCCIATCCG
                                                                            2940
50
                                                                            3000
      TTGAAGAACG GCGTTATCGA TAAAGAGCGG GTGCGAGAAG TATCGACTCG TCCCTACAAA
      CATCTTTCCG TCAATAAAAA TCTGGCTACT GCAAAACAGA TCTTTATTTC CTATTCCAAA
                                                                            3060
      GAGGATCAGA CTGAACTGGA GACCTGTCTG CAATTTTTCA AACCCTTGGA GAAGAATGGT
                                                                            3120
      CAGATCGAGA TCTACTATGA TAAGTTGACT AAGTTTGAAA CACCTATTCA CCCTGAAATA
                                                                            3180
      AGAAAGCGTA TTGTCGAAGC CGACTGTATA ATCGCTTTGA TCAGCCAACG CTATCTGGCC
                                                                            3240
55
                                                                            3300
      ACGGATTACA TCCTGGATCA TGAGTTGCCT GTATTTCGGG AGTATAACAA GACCATAGTG
      CCGATATTGA TCAAGCCTTG TACATTCGAA GACGATGAGT TCCTTCGGGA GAAATATTTT
                                                                            3360
      GCTCAGAAAG CTCAAATAAT CAATCTTGGA AAAGAGGGAA AAACCATTAA AGCTTATGAT
                                                                            3420
      AGTATTACGG CATCAGCCCA TCGTGATGAA AATTGGGTGG CAGTAGTCAG AGAGTTCAAA
                                                                            3480
      GAGAAGATAT TAAGAATAAC AAAACAGGAG GTAAATACAG ATGAA
                                                                            3525
60
      (2) INFORMATION FOR SEO ID NO: 254
           (i) SEQUENCE CHARACTERISTICS:
65
                (A) LENGTH: 687 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
70
          (ii) NOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
```

(iv) ANTI-SENSE: NO

75

| | (vi) ORIGINAL SOURCE: (A) ORGANISH: PORYPHYROHONAS GINGIVALIS | | | | | | | | | | | | |
|----|--|--|--|--|--|--|--|--|--|--|--|--|--|
| 5 | <pre>(ix) FEATURE: (A) NAHE/KEY: misc_feature (B) LOCATION 1687</pre> | | | | | | | | | | | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:254 | | | | | | | | | | | | |
| 10 | ATGATGAAAA AAGCATTTGT TTTCGTACTA CTGGTTTGCC TATTCTCCTC GTTCAGCAGT TCCGCCCCAAA CAACGACGAA CAGTAGCCGG AGTTATTTTA CAGGACGAAT CGAGAAGGTG AGTTTGAACT TAGGGGTCCC CCCCGTAAGC ACAGAGGTTT GGGGAATGAC CCATGATGCG AACGGTCTCC CTTTCGAAAT ACCTATCTCT TTCAGTCGTT TCAACAGCCA GGGAGATATA | 60 120 180 240 | | | | | | | | | | | |
| 15 | GCTACCACTT ATTACATAGC GAATAGCGAG GCAACTTTGA ATGAATGGTG CGACTATGCA CACCCGGGCG GCATCGTGAG GGTAGAAGGT CGTTTTTGGA AAATGACTTA CAACATACCA ACCTACAATG CAGTCTGCAC CCGGATTACA TTCGGAAAATC AAGAAATAGA AGGAACGATC GTCTTGATAC CCAAGCCCAA AGTCTCGCTG CCTCATGTGT CGGAATCGGT GCCTTGCATC CGAACCGAAG CCGGAGGGA ATTTATCCTT TGCGAAGAAG ACGACACCTT TGTGTCTCAC | 300 360 420 480 540 | | | | | | | | | | | |
| 20 | GATGGTAACG AAGTAACGAT AGGCGGTAAA CCTTTCTTGC TCAATACCAA CGTAAAGATT GTGGGGGGACG TATCTCAAAA GTATGCCGTG GGGGTAGGAG AAATTCGATT CCTGCAGATT TGTGCCCAAA CAGTATCACA ACAAAAA | 600 660 687 | | | | | | | | | | | |
| 25 | (2) INFORMATION FOR SEQ ID NO:255 | | | | | | | | | | | | |
| 30 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 684 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | | | | | | | | |
| | (ii) MOLECULE TYPE: DNA (genomic) | | | | | | | | | | | | |
| 25 | (iii) HYPOTHETICAL: NO | | | | | | | | | | | | |
| 35 | (iv) ANTI-SENSE: NO | | | | | | | | | | | | |
| | (vi) ORIGINAL SOURCE: (A) ORGANISM: PORYPHYROMONAS GINGIVALIS | | | | | | | | | | | | |
| 40 | (ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION 1684 | | | | | | | | | | | | |
| 45 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:255 | | | | | | | | | | | | |
| 50 | ATGAAAAAAG CATTTGTTTT CGTACTACTG GTTTGCCTAT TCTCCTCGTT CAGCAGTTCC GCCCAAACAA CGACGAACAG TAGCCGGAGT TATTTTACAG GACGAATCGA GAAGGTGAGT TTGAACTTAG GGGTCCCCC CGTAAGCACA GAGGTTTGGG GAATGACCCA TGATGCGAAC GGTCTCCCTT TCGAAATACC TATCTCTTTC AGTCGTTTCA ACAGCCAGGG AGATATAGCT ACCACTTATT ACATAGCGAA TAGCGAGGCA ACTTTGAATG AATGGTGCGA CTATGCACAC CCGGGCGGCA TCGTGAGGGT AGAAGGTCGT TTTTGGAAAA TGACTTACAA CATACCAACC | 60 120 180 240 300 360 | | | | | | | | | | | |
| 55 | TACAATGCAG TCTGCACCCG GATTACATTC GAAAATCAAG AAATAGAAGG AACGATCGTC TTGATACCCA AGCCCAAAGT CTCGCTGCCT CATGTGTCGG AATCGGTGCC TTGCATCCGA ACCGAAGCCG GGAGGAATT TATCCTTTGC GAAGAAGACG ACACCTTTGT GTCTCACGAT GGTAACGAAG TAACGATAAGC CGGTAAACCT TTCTTGCTCA ATACCAACGT AAAGATTGTG GGGGACGTAT CTCAAAAGTA TGCCGTGGGG GTAGGAGAAA TTCGATTCCT GCAGATTTGT GCCCAAACAG TATCACAACA AAAA | 420 480 540 600 660 684 | | | | | | | | | | | |
| 60 | | | | | | | | | | | | | |
| | (2) INFORMATION FOR SEQ ID NO:256 | | | | | | | | | | | | |
| 65 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1620 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: circular | | | | | | | | | | | | |
| 70 | (ii) MOLECULE TYPE: DNA (genomic) | | | | | | | | | | | | |
| | (iii) HYPOTHETICAL: NO | | | | | | | | | | | | |
| | (iv) ANTI-SENSE: NO | | | | | | | | | | | | |
| 75 | (vi) ORIGINAL SOURCE: | | | | | | | | | | | | |



```
(A) ORGANISH: PORTPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
 5
                 (B) LOCATION 1...1620
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256
      ATGAAGACAA AAGTTTTACG CAAATTCGTG GTGGCGGCTT TCGCCGTCGC AACCCTCTGT
10
      CCTCTCGCCC AAGCGCAGAC GATGGGAGGA GATGATGTCA AGGTGGTCCA GTACAATCAG
                                                                                120
      GAAAAACTGG TACAAACGAG GATGAGTGTG GCGGACAACG GATGGATCTA TGTAATGACC
                                                                                180
      CACAGTGGAT ACGACACCGG CAATAGCAAT GTGAAGATCT TCCGCTCCAA AGACCAAGGT
GCCACATACC AAAAGTTGAG GGATTGGGAT CCATCGGATG ATTATCAGTT TCAAGACTTC
                                                                                240
                                                                                300
      GATATOGTGG TAAOGGGTAA GAATGAATOO GACATOAAGA TTTGGTOGGT AGAGOTOATG
                                                                                360
15
      AATAAGCCCG GAGGATATAA GAGTAGAGTT GCGGTCTTCA GTCGCGATGC CAACGCGCAG
                                                                                420
      AATGCGAAAC TCGTGTATAA GGAAGACTTC TCCAATGTGC AGTTGTACGA TGTGGATATA
                                                                                480
      GCCTCCAACT ATCGTTCGCC TTCTTCTCTT AACAATGGTG GCAACCCTTT TGCTTTGGCT
                                                                                540
      TTCGCTTACA CCGGCTTCAA CAATACGCAC AAAATAAGTT TTGTGGACTA TGTGTTCTCT
                                                                                600
      CTGAATGGAG GGCAAAATTT CAATAAAAAC TTACTCTTCA GTCAAGATGG AGAGAAGAAA
                                                                                660
20
      ATTGACAAGG TGGATCTCTC ATTGGGTAGC ACCTCTGAAT CCATGGGTCA CAATGCCTGG
      CCGCTAATGG GTGTGGTATT CGAAATGAAT AAACAAGGGG GAAAAAGCGA TATCGGTTTC
                                                                                780
      TTGTCGAACT TTGTCGACAA TGATCCCGAA TTTCAGTGGT CAGGCCCTAT AAAAGTGAGT
                                                                                840
      GAAAGCGACA TGTCGTTCAG CCCCAAAATC CAAATGTTGC TGGACGAGGA TAACAATACG
                                                                                900
      ATCAATGGGG AGAGTTGCCA CAACTTCATG ATTACGTACA GCGATTATGA TTCTGAATAT
                                                                                960
25
      TCGGATTGGG ACATTCGGTA TGTATATCCC AAGAAATCGT TCAAGTATGA AAAAGGAAAA
                                                                               1020
      ACTCCGACTA TGGATGATCT GGTGGAAGCT TTCCTTACAG CTTCGTACCA GAGTGAGACC
                                                                               1080
      AACTCGGGGC TGGGGTATGA CAAGAACGCC AATCACTACC TGATTACATA TGCCAAAAAA
                                                                               1140
      GAAGAGACG GTACGAACAC GCTGAAATAC CGCTGGGCCA ATTATGACAA GATTCATAAC
                                                                               1200
      AAAGATTTGT GGAGCGACAC ATTTACGTAT ACATCATCTG CCAATGCTCT CTACACACCT
                                                                               1260
30
      CAAGTAGACA TCAATCCGAC CAAGGGTCTC GTGTGCTGGT CATGGGTGGA ATATCTGCCG
                                                                               1320
      GGCAAACGGA TCGTTTGGTC TGATACGCAG TGGACCCATG CCAACGGTGT AGAAGACATC
                                                                               1380
      GTANTGCAAG AAGGCAGCAT GAAGCTCTAC CCGAATCCGG CTCAAGAATA TGCTGTGATT
                                                                                1440
      ASCCTGCCGA CGGCAGCAAA CTGCAAGGCT GTTGTTTACG ATATGCAGGG CAGAGTAGTC
                                                                               1500
      GCTGAGGCTT CTTTCTCCGG CAACGAATAC AGGCTGAACG TGCAGCACTT GGCTAAGGGT
                                                                               1560
35
      ACGTACATAC TCAAGGTCGT ATCCGATACG GAGCGTTTCG TAGAGAAGCT CATCGTGGAA
                                                                               1620
      (2) INFORMATION FOR SEQ ID NO:257
40
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 2313 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
45
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
50
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
55
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...2313
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:257
60
      ATGCTGACGA TCCGAAACTT CCTCCTCTTT TGTTGTCTGT CGCTGATAGC GTTTGCTGCC
      GATGCACAAA GCTCTGTCTC TTCGGGTAGA CGACTGACAG AATATGTCAA TCCCTTTATC
                                                                                 120
      GGAACGSCCA ATTACGGTAC CACGAATCCG GGAGCAGTAT TGCCCAATGG GTTGATGAGC
      GTTACCCCTT TCAATGTCAG CGGATCGACA GAGAATCGCT TCGACAAAGA TTCGCGTTGG
TGGAGTGCGC CTTATTCGGC CGACAATAGT TACTGCATCG GTTTCAGCCA TGTGAATCTG
                                                                                 240
65
                                                                                 300
      AGTGGAGTAG GCTGTCCCGA ACTGAGTGGA ATACTGCTGA TGGCCACTTC CGGCACATTC
                                                                                 360
      GATCCTGATT ACTGCTGCTA TGGCTCTTCG CTCAGTCGAG AATATGCGCG CCCGGGAGAA
                                                                                 420
      TACAAGGCTG TATTGGACAA ATACGGTATA GATGCAGCCG TGACCGTAAC CGAGCGGACT
                                                                                 480
      GCTTTGACCG AATTTGCTTT TCCCGAAGGA GAAGGCCATA TCCTGCTGAA CCTGGGACAG
                                                                                 540
70
      GCCCTAAGCA ATGAATCGGG AGCCTCTGTT CGATTCTTAA ACGACTCCAC AGTCGTCGGC
                                                                                600
      AGCAGGCTGA TGGGGACGTT CTGCTACAAT CCGCAAGCAG TTTTTCGTCA GTATTTCGTA
                                                                                 660
      CTTCAGGTGA GTCGGCGACC GATCTCTGCC GGCTATTGGA AGAAGCAGCC TCCTATGACA
                                                                                 720
      GTGGAAGCCC AATGGGATTC GACTGCAGGG AAATATAAGC AGTACGACGG CTACAAGCGT
                                                                                 780
      GAGATGAGCG GTGATGACAT CGGTGTCCGA TTCTCGTTCA ACTGCGATCA GGGGGAAAAG
                                                                                840
75
      ATCTATGTAC GATCGGCCGT TTCATTCGTC AGCGAAGCCA ATGCGCTCTA TAATCTGGAA
                                                                                 900
```

Ĺ

```
960
       GCGGAGCAAG AAGAGGTGTT CAAAAGTGTC GGAGGGAATC CGGCCAAGGC TTTCTCCGCT
       ATACGCTCTC GCGCTATAGA GCGTTGGGAG GAAGCCCTCG GTACGGTGGA AGTGGAAGGA
                                                                                     1020
       GGCACACCGG ATGAAAAGAC GATATTCTAT ACCGCACTCT ATCACCTGCT GATACATCCG
                                                                                     1080
       AATATCCTAC AAGATGCCAA TGGAGAATAT CCTATGATGG GCAGTGGCAA AACGGGTAAT
                                                                                     1140
 5
       ACGGCTCACG ACCGCTACAC CGTGTTCTCT CTTTGGGACA CGTACCGCAA TGTACACCCG
                                                                                     1200
       CTGCTCTGCC TCCTCTATCC GGAGAAGCAG TTGGATATGG TACGGACACT GATCGACATG
                                                                                     1260
       TACCGAGAGA GCGGGTGGC'I GCCGAGATGG GAGCTGTACG GACAGGAGAC CCTGACGATG
                                                                                     1320
       GAGGGCGACC CCTCGCTTAT CGTCATCAAT GACACTTGGC AAAGGGGCCT TCGTGCTTTC
                                                                                     1380
       GATACGGCAA CGGCCTATGA AGCCATGAAA AAAAATGCTT CTTCGGCAGG AGCGACCCAT
                                                                                     1440
10
       CCGATCCGTC CTGACAACGA CGACTATCTC ACCCTCGGCT TCGTACCGCT TCGCGAACAG
                                                                                     1500
       TACGACAATT CCGTATCGCA TGCGCTGGAA TACTATCTGG CCGACTGGAA TCTGTCCCGG
       TTTGCCCACG CACTTGGGCA TAAAGAAGAC GCAGCTCTAT TCGGAAAACG CTCGTTGGGC
                                                                                     1620
       TACAGACACT ATTATAATAA GGAGTATGGT ATGCTGTGC CATTGCTGCC GGATGGATCA
                                                                                     1680
       TTCCTCACTC CTTTCGATCC CAAACAGGGT GAAAACTTCG AGCCTAATCC CGGTTTCCAC
                                                                                     1740
15
       GAGGGCAGTG CTTATAACTA TGCCTTTTTC GTTCCCCACG ATATACAAGG GCTTGCCCGG
                                                                                     1800
       CTGATGGGAG GAGCAAAGGT TTTTTCGGAA AGGTTGCAGA AAGTCTTCGA TGAAGGATAT TATGATCCGA CCAACGAGCC GGACATCGCC TATCCTTACC TCTTCTCCTA TTTCCCCAAG
                                                                                     1860
                                                                                     1920
       GAAGCATGGC GAACGCAGAA ATTGACCCGG GAGTTGATAG ACAAACATTT TTGCAATGCT CCTAACGGCT TGCCCGGTAA TGACGATGCC GGTACGATGA GTGCTTGGCT TGTCTATTCC
                                                                                     1980
                                                                                     2040
20
       ATGCTGGGAT TCTACCCTGA CTGTCCGGGC AGCCCCACCT ATACACTGAC CTCGCCGGTA
                                                                                     2100
       TTCCCCCGAG TTAGGATTCG GCTCAATCCG CAGTATTATC CTCAGGGGGA GTTGATCATT
                                                                                     2160
       ACGACCAATA CAGAGAATCA ACCGACAGAT TCCATTTACA TCCATACGGT TTCTCTTGGC
                                                                                     2220
       AATAAAACAC TTCCGCATGG AACAAGGCAT ATCAGCCATG CCGATTTGGT GCGCTGCGGT
                                                                                     2280
       CACCTCCGTT ACGAACTAAG CAATCGTCCT CGA
                                                                                     2313
25
       (2) INFORMATION FOR SEQ ID NO:258
             (i) SEQUENCE CHARACTERISTICS:
30
                  (A) LENGTH: 2328 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: double
                   (D) TOPOLOGY: circular
35
            (ii) MOLECULE TYPE: DNA (genomic)
           (iii) HYPOTHETICAL: NO
           (iv) ANTI-SENSE: NO
40
           (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: PORYPHYROHOMAS GINGIVALIS
           (ix) FEATURE:
45 .
                  (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...2328
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:258
50
       ATGTGCAAAA TAAGATTCAG CCTCTTGCAG GCTCTTGTCG TCTGCTTATT GTTCACCTCT
       TTTTCTCTCC AAGCTCAAGA GGAAGGTATT TGGAATACCC TCTTGGCTAT CCACAAGACG
                                                                                      120
       GAAAAAGCCG TAGAAACGCC CAAGAAAGTC TTTGCCGTAG CCAACGGAGT ACTTTACTCG
                                                                                      180
       GTGGGCAAAG AAGCTCCCCA TGAGGCAAAG ATCTTCGACC GTATCAGCGG ACTCAGCGAT
                                                                                      240
       ACATCGGTAA GCAGCATAGC CTACTCCGAG CAACTAAAAT CCTTGGTCAT ATACTATGCA
                                                                                      300
55
       TCAGGCAATA TCGACATCTT GGACGAAGCA GGCCGTGTGA CCAACGTACC TGCATTGAAA
                                                                                      360
       GACAATATCG ATCTGATAGA CAAAACGCTC AATCGCCTTT TGATCGTAGG CAACAGGGCT TATTTGGCAG GAGGATTCGG CCTCTCCGTT CTGGATGTCG CCGAAGCTCG CATACCGGCT
                                                                                      420
                                                                                      480
       ACCTACGCCA AGGGAACTAA GGTGACCGAT GTGGCTAAGT TGGACAATGA TCGCTTGCTG
                                                                                      540
       ATGCTGAAAG AAGGGCAGCT CTTCATCGGA AAAGAGACCG ATAACCTGCA AGATCCGGCC
                                                                                      600
60
       GCATGGACAG CCTTGTCTTT GAATTTGCCG ATGGGCTCGG TCACCGGTCT GGGCATTGTC
                                                                                      660
       GGGGAAGACA TCTGTTTCCT GCTCGCCGAT GGCCGTGTAT ATGTCGCTGC AAACCAATCG
                                                                                      720
       TTTGAGCCGG AGCTATTGCT CTCTTCCTCC GCCGATTCAC GACTGTATGT GACGGATCGT GGTCTGTTCA TCTGTGCCGA GAATCGAATT TATTTCATAG AAAAAGGTCG CAAAACGACA
                                                                                      780
                                                                                      840
       CAATTTCCTA TAGCCGACGT CCTTGGTGTC GGTGCCATGA ACGAAAGCAA TACGGCATAC ATAGCATTGG GAGAAGAAGG TTTGGCTTCA CTTCTTCTCG CAGAGGGAAG TACGGCCGAA
                                                                                      900
65
                                                                                      960
       GCCATGCCTG TAGCATTCGA CGGACCGGGG GACAATGATT TCTACGAGAT GCGGTTTAGT CACGGACGTC TGTATGCAGC CAGCGGACTC TGGGGAACAA ACCTGATGGG ACATGCCGGT
                                                                                     1020
                                                                                     1080
       ATGGTGAAGC TATACGACGG CAACCGATGG ACTAACTTCG ACAAGAAGAC CGTACAGGAA
                                                                                     1140
       CAGTTGGGCG GCGGATTCAG TTTCAATGAT GCTATCGATA TAGCTGTTTC CAACGGAGAC
                                                                                     1200
70
       CCCGATCACT TTTTTGTCGG TACATGGGGA AACGGTCTGT TCGAATTCAA GGATGGCAAA
                                                                                     1260
       GCGATAGCTC GCTATTCGGG AAACGAAACT GCTATCGCAG AATGTAATCC CGGAGATGCC
                                                                                     1320
       CGTGTGAAAG CGATTGCCTT TGACAATAAG GGCAACCTCT GGGGGACGCT CGGTGCCGTA
                                                                                     1380
       GGCAAGAACA TCTTCATGTA CGATCCGCAG AGTAGCACAT GGCATTCTTT CAGCTATCCG
                                                                                     1440
       GATGTAGCCA ATCTGGCCTC CTTCGGCAAT ATGATTATCC TACCCAACGG AGACAAATGG
                                                                                     1500
75
       GTAAATATCC TTCACCGTAG TGGCGGATCC ACGCGCAAAG GTGTCTTGAT CTTCAACGAT
                                                                                     1560
```

WO 99/29870 PCT/AU98/01023

```
CGGGGTACAC CGGAAACGAC TTCGGACGAC AGCCATCTTT ACGTCGAGCA GTTTGTCAAT
                                                                                1620
      CGCCTCGGGG CAGCCATAGG ACATAAGACT ATCTATGCAA TGGCCGTCGA TCATAACGGC
                                                                                1680
      TCTGTCTGGA TGGGATCGGA TATAGGCATT TTCGGCGTCT ACAATGCAGC CGGAGTATTG
                                                                                1740
      TCCTCGACTT CTACCCCTAT CGCTGTTCGG CCGGTCGGAG GAGAAGAACC CAATTTGTAC
                                                                                1800
 5
      TATGTGCTGG ACAAGGTGAC GGTGACAGAC ATCGTCGTGG ACAAACTCAA TCACAAATGG
                                                                                1860
      GTTGCCACCC AAGGGACAGG ACTCTATCTC CTTTCGGAAG ATTGCAGTAA GATCCTCGCG CAATTTACCG TAGAAAACAG CCCTTTGCTT TCTAACAACA TACTATCCCT GGCCTTAAAT
      GACGATAACG GACTGCTGTA CATCGGTACG GCGGACGGAC TGATGACGTT CCAAACGGGT
                                                                                2040
      ACGGGGAGTG GATCAGCTTC CGAACTGGAC GGCGTCTATG TATACCCCAA TCCGCTAAGG
                                                                                2100
10
      CCGGAATATC CCGATGGCGT CACCATTGCC GGACTGCAAG CCGGCTGTAG TGTCAAAATC
                                                                                2160
      ACCGATACCA CCGGCAGACT GCTATACCAG ACTGAGAGCG TAACCACCGA AGTCAAATGG
                                                                                2220
      AATGCTCGAG GTGCCGATGG CAATAGGGTA GCTTCGGGCG TATATGCCGT TGCAGTGTAC
                                                                                2289
      GATCCGGTAT CGAAAAAGTC CAAACTAATT CGCTTCGCAG TGATTCGC
                                                                                2328
15
      (2) INFORMATION FOR SEQ ID NO:259
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 3474 base pairs
20
                 (B) TYPE: nucleic acid
                 (C) STRAHDEDNESS: double
                 (D) TOPOLOGY: circular
           (ii) NOLECULE TYPE: DNA (genomic)
25
          (iii) HYPOTHETICAL: NO
           (iv) AUTI-SEHSE: NO
30
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
35
                 (B) LOCATION 1...3474
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:259
      ATGAAACGAA TACTTCCAAT AGTCGCATTC CTTTCTCTCT TCCTTGCCCT TGCTTTGCCT
40
      GCGAAAGCGC AACGAGCTAT GGGGAAGACG GCCGACCGTT CGCTAATGGC TTCGGGACAT
                                                                                 120
                                                                                 180
      TGGGTCAAGA TACGTGTCGA TGCAAGTGGA GTGTATCGCC TTACGGACGA ACAGCTCCGT
      GCCAATGGCT TCTCCGATCC GTCCAAGGTA GGTGTGTTCG GTTATGGTGG AGGGGTGCTT
                                                                                 240
      CCCGAAGATC TGAGCCGGAT CACGACAGAC GATTTGCCTC CGGTACCGGT ACTCCGTCAG
      GGCAATGCGC TGTATTTCTA TGCCGTGGGC CCGGTGACAT GGTTCTACAA TCCGGCCAAA
45
      ACCACCATGG AGCATACGGT GAATACATAC AGTACGCATG GCTACTACTT CCTGTCGGAT
                                                                                  420
      GCTGCCGGAG CACCTTTGCA GATGTCCCAA TATACGGGTG GAGGTGCGTC GGCCGAGGCT
                                                                                  480
      TTGATCGACT ACTACGATGA GCTGATGCTC CATGAACAGG AATTGTATTC GCCCAAAGAA TCGGGACGAG ATCTGTATGG CGAGTCTTTC AGTGCAGTCA ATACGCGTAC GGTCAAGTTC
                                                                                  540
                                                                                  600
      CCTTTGAGGG GCAACACCCG CTCGTCTGGC GAACTCGGTA CCGTATTCTC ATACATAGCC
                                                                                  660
50
      AAGGCCAGAT CGGCCGGTGG CGGCCGTGAG ATGTCGCTCT CGGCGAATGG CATTCTGATC
                                                                                 720
      TTCAGCGATC CTTTTTCCAT GACATCGAAT GAAGTCTCCA ATTCCTATTT GGCCGGCAAG
                                                                                 780
      AAGCGTCGTC TCTATCACAG TACGCCGATG AACAGCTTGG TCAATGAGTT GCGCTTGGAC
                                                                                 840
      GCGAACTATA GCATGACAGG AGATGCGGTC AATCTGGATT TCATAGAGGT GGCTACACAG
                                                                                  900
      AACGACCTCC GGTACGATGG CGCACCCATG CATATCAGGC GGTTTTCCAA TTTGCCCGTT
                                                                                 960
55
      TTGGGGGGCG AGTCCTGCCG GTTCGTTATC AGTGAGGTGC CGGAGTCTCT GGTGGTTTTG
      CAGGCCAATT CTTCCCTGAC AGCATCGCTT GTTCCCGTTA AGACTGTCGG GGATAAGACC
      ATTGAGTTCG TGGCTCCGCC GAAGGGTCAG GATCGTAGGA CTATCAATAC GTTTTATGCC
                                                                                1140
      GTGGACTTGT CACAGGCTTC TGCTCCGGAG ATCCTCGGAG CGGTACCCAA TCAAAACCTG
                                                                                1200
      CATGGAGAGG AAATCCCTGA TCTGATCATT GTCTCTACTC AGGCGCTCCT CCTTGAGGCT
                                                                                1260
      GATCGACTGG CCACCTATCG TAGAGAGAAA AACGGGCTGA AGGTTTTGGT CGTGTTGCAG
GAACAGGTGT TCAACGAGTT TTCGGGTGGA ACTCCCGATG CTACAGCATA CCGCCTCTTT
60
                                                                                1320
      GCCAAAATGT TCTACGACAG ATGGAAGGCA AATGCACCTG TGGGAGAGAC CTTCCCGATG
                                                                                1440
      CARATGOTTO TOTTOGGTGA TGGGGCTCAT GACAACAGGA AGGTCTCCGT AGCTTGGCAG
                                                                                1500
      AAACCGTATC TCCAACAAAC GGAGTTCTTG CTGACATTCC AAGCCGTCAA TTCGACGAAC
                                                                                1560
65
      GTAAACAGTT ATGTGACGGA TGATTACTTC GGCTTGCTGG ATGATCAGCC GGCCTCGGTC
                                                                                1620
      AATATCGGTT GGCGCAATTA TAATATGGCT GTAGGGCGAT TCCCCGTACG TACTCCGGCC
                                                                                1680
      GAAGCTCGCA TCGCAGTGGA CAAGACCATC CGATATGAGG AGGATCGAGA GAGTGGTGCC
                                                                                1740
      TGGCGTATTC GTGCCTGTTT TGCGGCAGAC AACGGGGACA AGCACGCAAC CGAGACTTCC
                                                                                1800
       CGTTTGATCG ATACCGTCAA GCGTTATGCT CCTGCCATCA TGCCGGTACG CGCCTTTCAG
                                                                                1860
70
      GACGTATATC CGCATGTCAT CGAGAACGGG TTGCACAGCA TTCCGGGTGC AAAGAAAAAG
      ATGCTGGAAA CCCTTCAGTC GGGTATTATC CTGCTTAATT ATGCTGGTCA TGGCGGTCCT
                                                                                 1980
      GCCGGATGGT CGGACGAGCA TTTGCTGACG CTCAACGATA TACACAAATT CAATTATAAG
                                                                                2040
      CATATGCCCA TTTGGATTAC TGCCACGTGC GACTTTGCCA ACTATGACAG TCAGACGACC
                                                                                2100
      TCGGCAGGGG AGGAGGTTTT CCTCCATGAG AAGAGTGGCA CTCCGATCAT GTTCTCGACT
                                                                                2160
75
      ACGCGTGTCG TTTACAATAC GCAGAATGAG AAGATCAATG GTTTTATGCT TCGGCGTATG
                                                                                2220
```

PCT/AU98/01023

```
TTCGAGAAAG CTAAGGATGG GCGTTATCGT ACGATGGGCG AGATTATCCG ATCGGCCAAA
                                                                             2280
       CAGGGGATGC TCAGTACTGT TTTCCCCGAT TCGATCAACC AGTTGAGTTT CTTTCTGATG
                                                                             2340
       GGTGATCCGT CCGTGCGTAT GAATCTTCCT ACCCACAAAG TGCAATTGAC CGCAATCAAC
                                                                             2400
      GGGCAGGATC CCGAAGGGCA GTATGGAACT ATTATGCTCA AGTCTTTGGA ACGGGTAGCT
                                                                             2460
 5
      CTGAAGGGTA AGGTAACCGA TGAAAAGGGG ACATTCGACG AGACATTCAG TGGCAAGGTT
                                                                             2520
      TTCCTGACCG TCTTCGATGG CAGAAAGAAA ATGACAGCTT TGGAAGAGGA GGGAAACGAT
                                                                             2580
      CTCTCTCTG TATATTATGA CTATCCTAAC GTGATGTATG CCGGTATTGC CGAGGTGAAA
                                                                             2640
      GACGGACTCT TCGAAACTTC GTTTATCGTA CCCAAGGATG TGAACTATTC CGAGCACGAA
                                                                             2700
       GGCCGGATCA ATCTTTATGC TTATAACGAG AGCACAAAGG CGGAAGCCAT GGGGGTAGAC
                                                                             2760
10
       TTCTCCATCA GAGTCCAACC GGGTATTCCT GATGAGGTAA CGGAAGATAA TACACCGCCT
                                                                             2820
       GAAATCATAA GCTGCTTCCT CAATGACAGT ACATTCCGAT CGGGAGATGA GGTTAATCCT
                                                                             2880
      ACTCCTCTGT TTATGGCCGA AGTATTCGAC TTGAATGGAA TCAATATCAC GGGTAGCGGA
                                                                             2940
      GTAGGGCATG ATATTACGCT TTGTATCGAT GGCCGTGCCG ACCTGACCTA CAACCTCAAT
                                                                             3000
      GCATATTTCA CAAGTTCGGC TACGGATGCA GGTGTGGGCA CTATTCTCTT CATGATACCG
                                                                             3060
15
      GCTTTGGCCG AAGGAGATCA TACTGCCCGA CTGACGGTTT GGGACATTTT CAATAATGCC
                                                                             3120
      GTCCATCATG ACTTTTCATT CAGAGTGGTA GATGGCATTG CTCCGGATGT GGCTGATGTG
                                                                             3180
      ATTCTATTCC CGAATCCGGT ACGCGAGAGT GCTACGTTCC GAATCTTCCA CAATCGCCCC
GGAAGCGATT TGAACGTGGC CGTGGAGATC TATGACTTCA CCGGTCGTCT TGTGAACAGT
                                                                             3240
                                                                             3300
      TTGCCAGTCA AGACCTATTC GTCTTCCTAC GGAGAACCTA TAGAGATCAA GTGGGATCTG
                                                                             3360
20
      ACCTCCAAAT ACGGAGTGAA GATCGGAAAC GGATTCTACC TCTATCGTTG TGTGGTGAAC
                                                                             3420
      TCTCCCGGAG GACAGACGCC CTCCATGGCC AAGAAAATGA TCGTGGTAGG ACAA
                                                                             3474
       (2) INFORMATION FOR SEQ ID NO: 260
25
            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 2883 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
30
                 (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HYPOTHETICAL: NO
35
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
40
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                (B) LOCATION 1...2883
45
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:260
      ATGAAGAAAC TTTTCCCATT ATTATTACTC ATATTGTCGA TTTTGGTCGG ATGTGGAAAA
      AAAGAAAAC ACTCTGTAAC TGAAATCGCC CGAGAGAAAA AGCGTATTAC TGCATTGCTG
      TACGAAAAAG AACTCCCCAC TGATTCTGTT AAACAGCTTT ACGAGAACAG CGTACAGAAC
                                                                              180
50
      AAGAACCTTG TGGGACAAAT GTTATTCGCG ATCGAGGTCG GCAAACGGAT GCGTAATATG
      TCCCAATACA CAGATGCGAT GCTATATCAC CAAAAAGGGT TGAACGCTGC ATTGAATCTA
                                                                              300
      AGGGACACCA TCGTAGCCGC ACAAGCATGG AATCATCTCG GAACGGATAG CCGTCGTATC
                                                                              360
      GGTGCTTTGG CAGAAGCTTC GGATTATCAT TACAAGGCTC TTTCCTTGAT AGAATCTTTT
                                                                              420
      AGCGGAAACC AGAATAGGCC TGCGATCAAA GCCAGATCGG CGGCCCTGAA CGGCATCGGC
                                                                              480
55
      AACATCAATC TTGAGTTAGG ATACCATGAT GAGGCCGAAA AGAATTTCCT GAAAGCACTG
                                                                              540
      CAAGGTGAGA AAGAACTCGA CAGTCCTCTC GGGCAAGCTA TCAACTATGC GAACCTCGGA
                                                                              600
       CGTATCTATC GACAACGCAA AGAATACGAC AAGGCTCGTA CCTACTTCCT CTTGTCTCTG
                                                                              660
      GAGCAGAACA ATATGGCAGA GAATCTGATG GGTATCGGAC TCTGTAGCAT CAATCTCGGA
                                                                              720
      GAAGTAGACG AAGAAAAAGG GGATTATCAA AAGGCTTTGC AAGAGTATGC CACGGCATAC
                                                                              780
60
      AAACTGATGG AACAGTTGTC CGATCGATGG CACTGGCTGA ATTCCTGTAT CCCGATGGCA
                                                                              840
      900
      GGGACTGCGA AAGAAATTAA TTCGACTTCA CATCTGATAG AAATATACAA TCTTCAATAC
                                                                              960
      GAGAATCTCG AGCGTAAAAA AGAATACAAA CAAGCCCTCG AAGCATTCTG TCTGAGCAAG
                                                                             1020
      ACGTTGAGCG ACAGCATGTC CATTGCGCAC AAGGTCAGCA GCATACAAGA AACGCGATTC
                                                                             1080
65
      AACTACGAAC GAAACAAGTC CCAAAAAGAG CTTGAAGAAA TACAGCAAGT AAGCAAGGCA
                                                                             1140
      AAACAAGAGA AATCGAAGTT TATCCTCTTG AGCACTCTTT TTGCCCTTTT CATCTCGATT
                                                                             1200
      CTTTTGATTT CTGTTCTGAC ATATGCATAC CGTCAGGGCA AGAAGCATAA CAAGCTGATC
                                                                             1260
      AAAGAGACGG ATAAACTTCG CTCCGGCTTT TTCACCGGTA TTACACACGA ATTTCGTACG
                                                                             1320
      CCTATCACCG TCATACAAGG TTTGAATGAG AAAATGAGTT CAAGTCCTGA TCTCCAAGCA
                                                                             1380
70
      TCGGACAGAA CCGAGCTGCA CAAGATAATA GACAGACAGA GTAGCCATAT GCTGAATTTG
                                                                             1440
      GTGAACCAGC TGTTGGATAT TTGCAAGATC AGAAGCGGAG TATCCACGCC CGAATGGCGC
                                                                             1500
      AATGGCGACA TCGTCTCCTT CGTACAGATT CTCATCGATT CGTTTGCACC ATACGCACAG
                                                                             1560
      GCTCAAGACA TAACCTTGGA GCTACAACCC GAGAGCAAAC CTATTGTCGT GGACTTCGTC
                                                                             1620
      CCCTCCTACT TGCAAAAAAT CATATCCAAT CTTTTGTCCA ATGCCATCAA GTATTCTTTA
                                                                             1680
75
      GCCGGAGGGA GAGTGGTCAT ATCTCTGGCA AAAACCAAGA ATGAAAAAAA TCTGATCATA
                                                                             1740
```

```
CGCGTTGCAG ACAATGGCAT AGGAATAGAT AAAACTGATC AGGCTCATAT CTTCGACATC
                                                                               1800
      TTCTATCGAG GACAGTCCGC TACCGAAAAG CATGGATCAG GCGTCGGACT CTCGTTTACC
      AATATACTGG TCGAAAACCT TCGAGGTACG ATCAAAGTGG AAAGCCAGCC GGGGAAAGGA
                                                                               1920
      AGTGCCTTCA CCATCAGTAT TCCTACACAA AACCAGTCCT CTTCGGCAGA GATTCTTCCT
 5
      TGGCTACCCT CCTCCGATGA CATTGTCATG CCTGTCCACA TCGCGCCCGA TGACTCACCG
                                                                               2040
      ACATCTCCGA TGGTAGCAGC TCTGAATCAT CGCTTCGAGG ACGAACGTCC GACCATACTG
                                                                               2100
      CTCGTCGAGG ACAATAAGGA TATCAACCTG CTCGTCAAAC TACTCCTTTG CGATCGCTAC AATGTGCTAT CCGCCGCAAA CGGAAAAGAG GGTATAGCCC TCGCTACCGA GCATATTCCC
                                                                               2160
                                                                               2220
      GACATTATCA TTACGGATAT TATGATGCCG ATAATGGATG GGATAGAAAT GACAATCCGG
                                                                               2280
10
      ATGAAGCAAT CGCCTCTGCT CTGTCACATT CCCATTGTCG CTTTGACGGC CAAGAGTACC
                                                                               2340
      GAACAGGACA GATTGGAAGG AATCAAAAGC GGTGTAGTCT CTTATCTATG CAAGCCATTC
                                                                               2400
      TCTCCGGAGG AGCTTTTGAT GCGGATCGAG CAGCTTCTGA AAGACCGTGA GTTGCTCAAG
                                                                               2460
      AAGTTCTATA TGCAAAAACT CATGCTGGAT CGGAAGCCGG AGGAGGAGCC TCAACCGATA
                                                                               2520
      GATGACAGCA GTATGCAGTT TCTCCTTGCT GCCAAAGATG CAGTGTCCGG TGGAATCAAA
                                                                               2580
15
      CAAAATCCGG ATTTTTCCGC TCAAGACTTG GCCGAAAAAA TGTGCATGAG TCCATCCCAA
                                                                               2640
      CTCAACAGAA AGCTCACGAG TGTCGTAGGT TGCTCCACCA TCGGCTACAT ACAGCAGATC
                                                                               2700
      AAGATAAAAT TGGCCTGCAA GCTCCTTGCC GATGAGAGCA AAAACATCTC CGACATTAGC
                                                                               2760
      ATTGAGGCAG GCTTTTCGGA TCCGGCTTAC TTCTCTCGCA CCTTCAAACG CTACATGAAC
                                                                               2820
      TGCTCTCCCT CCCAATATCG GCAAAAACTC CTTGCCATGC CGGGGAGCGA CAAGGAGACA
                                                                               2880
20
      GTT
                                                                               2883
      (2) INFORMATION FOR SEO ID NO: 261
25
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1668 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                 (D) TOPOLOGY: circular
30
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
35
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: PORYPHYROMONAS GINGIVALIS
40
          (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...1668
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:261
45
      ATGAAAAGC TTTTACAGGC TAAAGCCTTG ATTCTGGCAT TGGGACTCTT CCAACTGCCC
                                                                                 60
      GCAATCGCCC AAACGCAAAT GCAAGCAGAC CGAACAAACG GTCAATTTGC AACAGAAGAG
                                                                                120
      ATGCAACGAG CATTCCAGGA AACGAATCCC CCTGCAGGTC CTGTGCGTGC TATCGCTGAG
                                                                                180
      TACGAACGCT CTGCAGCCGT TTTGGTACGC TACCCGTTCG GTATCCCGAT GGAATTGATC
                                                                                240
50
      AAAGAGCTGG CCAAGAACGA CAAGGTGATT ACCATTGTGG CGAGTGAAAG CCAAAAAAAAC
                                                                                300
      ACCGTTATAA CCCAGTACAC CCAAAGCGGT GTGAATCTCT CTAATTGCGA TTTCATCATT
                                                                                360
      GCGAAAACTG ACTCTTACTG GACACGCGAC TATACCGGTT GGTTCGCAAT GTACGATACG
                                                                                420
      AACAAAGTAG GTCTCGTGGA CTTTATTTAT AACCGCCCTC GTCCTAACGA TGATGAATTC
                                                                                480
      CCCAAATACG AAGCACAATA TCTGGGCATC GAGATGTTCG GGATGAAGCT CAAGCAGACC
55
      GGTGGCAACT ACATGACGGA CGGATATGGA TCCGCTGTGC AGTCACATAT CGCATATACG
                                                                                600
      GAGAACTCCT CTCTGTCTCA AGCTCAAGTA AATCAAAAGA TGAAAGACTA TCTCGGCATC
                                                                                660
      ACACATCATG ATGTGGTACA AGATCCGAAC GGCGAATATA TCAACCATGT GGACTGTTGG
                                                                                720
      GGCAAGTATT TGGCACCGAA CAAAATCCTC ATCAGGAAAG TGCCTGACAA TCACCCTCAG
                                                                                780
      CACCAAGCCC TGGAAGATAT GGCAGCCTAC TTCGCAGCAC AGACCTGCGC ATGGGGAACG
                                                                                840
60
      AAGTACGAGG TATATCGCGC TTTGGCCACC AATGAACAAC CGTACACGAA CTCTCTGATT
                                                                                900
      CTGAACAACA GGGTATTTGT TCCTGTCAAT GGCCCCGCCT CCGTGGACAA CGATGCTCTG
AACGTCTATA AGACGGCAAT GCCCGGTTAC GAAATTATAG GTGTCAAAGG GGCTTCAGGA
                                                                                960
                                                                               1020
      ACACCTTGGT TAGGAACAGA TGCCCTGCAT TGTCGTACTC ACGAGGTAGC GGATAAGGGC
                                                                               1080
      TATCTCTATA TCAAGCACTA CCCGATACTG GGCGAACAGG CAGGCCCTGA TTATAAGATC
                                                                               1140
65
      GAAGCAGATG TCGTCTCATG CGCCAATGCT ACTATCTCGC CGGTACAATG TTACTATCGT
                                                                               1200
      ATCAATGGTT CCGGTAGCTT TAAGGCTGCT GATATGACGA TGGAATCAAC AGGTCACTAT
                                                                               1260
      ACTTATAGCT TTACAGGTCT TAACAAGAAT GATAAGGTAG AATACTATAT CTCTGCCGCT
      GACAATAGTG GTCGCAAAGA GACTTATCCC TTTATCGGCG AACCTGATCC TTTCAAGTTT
                                                                               1380
      ACGTGTATGA ACGAAACCAA TACATGTACT GTGACCGGAG CTGCCAAAGC TCTTCGTGCA
                                                                               1440
70
      TGGTTCAACG CCGGTCGTTC AGAACTGGCT GTTTCGGTAA GTTTGAATAT TGCCGGCACA
                                                                               1500
      TATCGGATAA AGCTTTATAA CACCGCAGGA GAAGAAGTCG CTGCAATGAC CAAGGAATTA
                                                                               1560
      GTAGCAGGGA CGAGTGTCTT CAGTATGGAT GTGTATTCTC AGGCTCCGGG CACATATGTT
                                                                               1620
      CTGGTTGTTG AAGGAAATGG AATCCGTGAG ACAATGAAAA TTCTCAAA
                                                                               1668
75
```

```
(2) INFORMATION FOR SEQ ID NO: 262
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1284 base pairs
 5
                (B) TYPE: nucleic acid
                 (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
10
         (iii) HYPOTHETICAL: NO
          (iv) ANTI-SEUSE: NO
15
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
20
                (B) LOCATION 1...1284
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262
      ATGAAACTTT CATCTAAGAA AATCTTAGCA ATCATTGCAT TGCTGACGAT GGGACATGCT
25
      GTGCAGGCAC AGTTTGTTCC GGCTCCCACC ACAGGGATTC GCATGTCTGT CACTACAACC
                                                                             120
      AAGGCCGTAG GCGAAAAAAT CGAATTGTTG GTTCATTCCA TAGAGAAGAA AGGCATCTGG
                                                                             180
      ATCGATCTCA ATGGGGATGC CACTTACCAA CAAGGAGAG AAATAACCGT ATTCGATGAG
                                                                              240
      GCATACCACG AATACACGAT CGGGACGCAA ACCCTCACTA TCTATGGTAA TACGACCCGA
                                                                              300
      TTGGGCTGTC GATCTACCGG TGCAACGGCT GTCGATGTAA CGAAAAACCC TAATCTGACC
                                                                              360
30
      TATCTCGCAT GCCCGAAAAA TAATCTGAAA TCATTGGACT TGACGCAAAA CCCAAAGCTG
                                                                              420
      CTGCGAGTTT GGTGCGACTC TAACGAAATA GAAAGTTTGG ACCTGAGTGG CAATCCGGCT
                                                                              480
      TTGATCATCC TCGGCTGTGA CAGGAATAAG CTGACTGAGC TGAAGACCGA TAACAACCCC
                                                                              540
      AAGTTGGCCT CTCTTTGGTG TTCTGATAAT AACCTGACGG AGTTGGAACT CAGTGCCAAT
                                                                              600
      CCTCGTCTCA ATGATCTTTG GTGCTTCGGT AATCGGATCA CGAAACTCGA TCTGAGTGCC
                                                                              660
35
      AATCCTCTAT TGGTAACACT TTGGTGCAGT GACAATGAGC TTTCGACCTT GGATCTTTCC
                                                                              720
      AAGAATTCGG ACGTTGCTTA CCTTTGGTGT TCATCGAACA AACTTACATC CTTGAATCTG
                                                                              780
      TCGGGGGTGA AGGGACTGAG TGTTTTGGTT TGTCATTCCA ATCAGATCGC AGGTGAAGAA
                                                                              840
      ATGACGAAAG TGGTGAATGC TTTGCCCACA CTATCTCCCG GCGCAGGCGC TCAGAGCAAG
                                                                              900
      TTCGTCGTTG TAGACCTCAA GGACACTGAT GAGAAGAATA TCTGTACCGT AAAGGATGTG
                                                                              960
40
      GAAAAAGCTA AAAGTAAGAA CTGGCGAGTA TTTGACTTCA ACGGTGATTC TGACAATATG
                                                                             1020
      CTTCCATACG AAGGAAGTCC GACATCGAAC TTGGCAGTAG ATGCTCCCAC TGTCAGGATA
                                                                             1080
      TATCCCAATC CGGTAGGAAG ATATGCGCTC GTCGAGATCC CCGAGTCTCT TTTAGGGCAG
                                                                             1140
      GAAGCTGCTT TATACGATAT GAATGGGGTA AAAGTCTATA GTTTCGCGGT AGAGTCTCTT
                                                                             1200
      CGTCAGAACA TTGACCTGAC ACATCTTCCC GACGGCACTT ATTTCTTCCG TCTCGATAAC
                                                                             1260
45
      TATACCACTA AGCTCATCAA ACAG
                                                                             1284
      (2) INFORMATION FOR SEC ID NO: 263
50
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 930 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: circular
55
          (ii) MOLECULE TYPE: DNA (genomic)
         (iii) HYPOTHETICAL: NO
60
          (iv) ANTI-SENSE: NO
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
65
          (ix) FEATURE:
                (A) HAME/KEY: misc feature
                (B) LOCATION 1...930
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263
70
      ATGAGAAAAA CAATAATTTT CTGCTTGTTG CTCGCCCTAT TTGGCTGTTC TTGGGCACAA
                                                                               60
      GAAAGAGTCG ATGAAAAAGT ATTCTCCGCA GGAACAAGTA TTTTTAGGGG CATCCTTGAA
                                                                              120
      AAGGTGAAAG CACCGCTTAT GTATGGAGAT CGTGAGGTAT GGGGTATGGC TCGTGCGAGC
                                                                             180
      GAGGATTTCT TTTTTATACT TCCCGTTACG GATGACCTCA CTCCCGTGCT TTTCTATAAC
                                                                              240
75
      CGTCTTACAA ACGAACCCTG CTTTGTGTCA GACCAAGGAA TAACTGAGTA TTTCAAATTC
```

```
GCTCAAGAAG GTGATTACAT TGAAGTCGAA GGAAGCTCTG TATTCATGGC GAATCTTTTG
                                                                              360
      TACTATCGTT TTTTCCCGAC AAGAATTACC TCCTATAATG CTCCCATTGA AGGTGTTGTG
                                                                              420
      AGCAAGACGG GAAATCCTGC TTTTACAATC CCGATGCTCC CGGGGGTTTC TGATTGCATA
                                                                              480
      GAAATCTCAA ACAACCGCAA AGTCTTTCTG ACCAATCAAT TAGGGGTTGT AAACATCACT
                                                                              540
      GACGGGATGG AACCTCCGAT TATTGCCGGA GTCTCTGCTT CCTATGGATC TTCCGTCCGG
                                                                              600
      GTGTATGGTC ATGTCTCACA GCGGTGGGAC ATCATAGGCC ATTGCTATTT GGATATCTAC
                                                                              660
      CCAACCAATT GCTATCCGCT CAGCACGAAA CCCGTTGCAG GAGACGATGA GGTTTTTGTC
                                                                              720
      AAACAACAAG GCAGGCAAAT AGAGATCGAT AGCAACAGCC CCATAGTCCA AGTGGTCGTA
                                                                              780
      TACGATCTTG AGGGGAAAAG TGTTTTTCGC AAAAGAATGA CCGAAAACGC TTATACCCTA
                                                                              840
10
      TCCTTTAGAG CACCCATGCT CGGCTTTATG ACCATCATGA TCGAAACACA AAATTCGATT
                                                                              900
      ATCAATAAAA AACTTAATGT TACACAGCTA
                                                                              930
      (2) INFORMATION FOR SEQ ID NO: 264
15
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 1215 base pairs
                 (B) TYPE: nucleic acid
                 (C) STRAHDEDHESS: double
20
                 (D) TOPOLOGY: circular
          (ii) MOLECULE TYPE: DNA (genomic)
          (iii) HTPOTHETICAL: NO
25
          (iv) AHTI-SEHSE: NO
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: PORYPHYROMONAS GINGIVALIS
30
          (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...1215
35
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:264
      ATGAAAAAA CAACCATTAT TTCTTTGATT GTCTTCGGTG CTTTCTTTGC AGCCGTGGGC
      CAAACCAAGG ACAATTCTTC TTACAAACCT TTTTCGAAAG AAGATATTGC CGGAGGAGTT
                                                                              120
      TACTCTCTCC CGACTCAAAA TCGTGCGCAG AAGGACAATG CCGAGTGGCT TCTTACAGCG
                                                                              180
40
      ACCGTCTCCA CAAACCAGTC TGCAGATACT CACTTTATCT TCGATGAGAA CAACCGCTAT
                                                                              240
      ATCGCTCGTG ACATAAAAGC CAATGGGGTA AGAAAATCCA CGGACTCCAT TTACTACGAT
                                                                              300
      GCCAACGGGC GAATATCGCA TGTGGATCTT TATATCTCGT TCAGTGGCGG AGAGCCTGCA
                                                                              360
      CTCGACACCC GATTCAAGTA CACCTATGAT GACGAGGGAA AGATGACCGT GAGGGAAGTA
                                                                              420
      TTCATGCTGG TAATGGATCC GAATACACCT ATCTCACGCT TGGAATATCA TTATGATGCA
                                                                              480
      CAGGGCAGAC TGACCCACTG GATTTCTTTT GCTTTCGGGG CAGAATCCCA AAAGAATACG
45
                                                                              540
      TATCACTATA ATGAAAAAGG TCTGTTGGTC AGCGAAGTGC TGAGCAATGC AATGGGGACA
                                                                              600
      ACCTATTCAG ACACCGGCAA AACGGAATAC AGCTATGACG ATGCAGATAA TATGGTGAAG
                                                                              660
      GCCGAGTACT TCGTCGTCCA GCAAGGAAAG GCATGGCAAG TACTCAAAAG AGAGGAATAC
                                                                              720
      ACCTATGAGG ACAATATCTG CATACAATAT TTGGCTATTA ACGGTACCGA CACAAAGGTG
                                                                              780
50
      TACAAGCGAG ACATCGAGAG CGATAAGTCC ATCTCCGCAA ATGTCATTGA CATTCCGTCA
                                                                              840
      ATGCCGGAAC AGACCTGGCC TAATATGTAC GGATTCAACG CAAAGCGACT GAAAGAGACT
                                                                              900
      TATTCCTCCT ACGAAGGAGA TGTGGCTACT CCTATATTCG ACTATATCTA TACGTACAAG
                                                                              960
      GCTCTTACCT CAATGCCAAC ACCTTCGACA GAAGCTCAGG TAGCAGTCTA TCTCAATCCG
                                                                             1020
      TCAACGGACC GGTTAGTGAT TCTGGCCAAC GGCATCACAC ATCTGAGCAT GTACGACTTG
55
      CAGGGTAAGC TTATCCGTGA TTGTGCCTTG AGCGGCGATA AGGTGGAAAT GGGTGTCGGA
                                                                             1140
      TCTTTGACCA AAGGGACATA CCTGCTTAAA GTGAATACGG ATCAGGGAGC CTTTGTGAGA
                                                                             1200
      AAAGTCGTGA TTCGA
60
      (2) INFORMATION FOR SEQ ID NO: 265
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 454 amino acids
                 (B) TYPE: amino acid
65
                (D) TOPOLOGY: linear
          (ii) MOLECULE TYPE: protein
         (iii) HYPOTHETICAL: YES
70
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
          (ix) FEATURE:
75
                (A) NAME/KEY: misc_feature
```

(B) LOCATION 1...454

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265
```

```
5
       Phe Cys Val Met Ala Lys Val Ile Lys Thr Lys Lys Gly Leu Ala Leu
       Asn Leu Lys Gly Lys Pro Leu Pro Glu Met Leu Ala Glu Pro Ala Gln
20 25 30
       Ser Pro Thr Tyr Ala Val Val Pro Asp Asp Phe Glu Gly Val Ile Pro
10
                                        40
       Lys Val Thr Ala Arg Pro Gly Asp Lys Val Arg Ala Gly Ser Ala Leu
50 60
       Net His His Lys Ala Tyr Pro Glu Met Lys Phe Thr Ser Pro Val Ser
65 70 75 80
       Gly Glu Val Ile Ala Val Asn Arg Gly Ala Lys Arg Lys Val Leu Ser
85 90 95
15
       Ile Glu Val Lys Pro Asp Gly Leu Asn Glu Tyr Glu Ser Phe Pro Val
100 105 110
       Gly Asp Pro Ser Ala Leu Ser Ala Glu Gln Ile Lys Glu Leu Leu
115 120 125
       115 120 125
Ser Ser Gly Met Trp Gly Phe Ile Lys Gln Arg Pro Tyr Asp Ile Val
130 135 140
20
       Ala Thr Pro Asp Ile Ala Pro Arg Asp Ile Tyr Île Thr Ala Asn Phe
145 150 160
25
       The Ala Pro Leu Ala Pro Asp Phe Asp Phe Ile Val Arg Gly Glu Glu 165 170 175
       Arg Ala Leu Gin Thr Ala Ile Asp Ala Leu Ala Lys Leu Thr Thr Gly
180 185 190
       Lys Val Tyr Val Gly Leu Lys Pro Gly Ser Ser Leu Gly Leu His Asn
195 200 205
30
       Ala Glu Ile Val Glu Val His Gly Pro His Pro Ala Gly Asn Val Gly 210 215 220
       Val Leu Ile Asn His Thr Lys Pro Ile Asn Arg Gly Glu Thr Val Trp
225 230 235 240
35
       Thr Leu Lys Ala Thr Asp Leu Ile Val Ile Gly Arg Phe Leu Leu Thr 245 250 255
       Gly Lys Ala Asp Phe Thr Arg Met Ile Ala Net Thr Gly Ser Asp Ala
260 265 270
       Ala Ala His Gly Tyr Val Arg Ile Met Pro Gly Cys Asn Val Phe Ala
275 280 285
40
       Ser Phe Pro Gly Arg Leu Thr Ile Lys Glu Ser His Glu Arg Val Ile
290 295 300
       Asp Gly Asn Val Leu Thr Gly Lys Lys Leu Cys Glu Lys Glu Pro Phe 305 310 315 320
       Leu Ser Ala Arg Cys Asp Gln Ile Thr Val Ile Pro Glu Gly Asp Asp
325 330 335
45
       Val Asp Glu Leu Phe Gly Trp Ala Ala Pro Arg Leu Asp Gln Tyr Ser
340 345 350
       Het Ser Arg Ala Tyr Phe Ser Trp Leu Gln Gly Lys Asn Lys Glu Tyr
355 360 365
50
       Val Leu Asp Ala Arg Ile Lys Gly Gly Glu Arg Ala Met Ile Het Ser
370 375 380
       Asn Glu Tyr Asp Arg Val Phe Pro Net Asp Ile Tyr Pro Glu Tyr Leu
385 390 395 400
55
       Leu Lys Ala Ile Ile Ala Phe Asp Ile Asp Lys Met Glu Asp Leu Gly
405 410 415
       Ile Tyr Glu Val Ala Pro Glu Asp Phe Ala Thr Cys Glu Phe Val Asp 420 425 430

Thr Ser Lys Ile Glu Leu Gln Arg Ile Val Arg Glu Gly Leu Asp Met
60
                435
       Leu Tyr Lys Glu Met Asn
        (2) INFORMATION FOR SEQ ID NO: 266
65
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 201 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
70
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
```

75

(vi) ORIGINAL SOURCE:

```
(A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
 5
                    (B) LOCATION 1...201
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:266
       Glu Leu Ser Lys Cys Tyr Met Asp Lys Val Ser Tyr Ala Leu Gly Leu 1 5 10 15
10
       Ser Ile Gly Asn Asn Phe Lys Ser Ser Gly Ile Asp Ser Val Val Het 20 25 30
       Asp Asp Phe Net Gln Gly Leu Ser Asp Val Leu Glu Glu Lys Ala Pro
35 40 45
       Gln Leu Ser Tyr Asp Glu Ala Lys Arg Glu Ile Glu Ala Tyr Phe Het
50 55 60
15
       Asp Leu Gln Gln Lys Ala Val Lys Leu Asn Lys Glu Ala Gly Glu 65 70 75
        Phe Leu Lys Ile Asn Ala His Lys Glu Gly Val Thr Thr Leu Pro Ser
85 90
20
        Gly Leu Gln Tyr Glu Val Ile Lys Met Gly Glu Gly Pro Lys Pro Thr
100 105 110
        Leu Ser Asp Thr Val Thr Cys His Tyr His Gly Thr Leu Ile Asn Gly
115 120 125

    Ile Val
    Phe Asp
    Ser
    Ser
    Met Asp
    Arg
    Gly
    Glu
    Pro Ala
    Ser
    Phe Pro 130

    Leu Arg
    Gly
    Val
    Ile
    Ala
    Gly
    Trp
    Thr Glu
    Ile
    Leu Gln
    Leu Met
    Pro 140

    145
    150
    150
    155
    160

25
        Val Gly Ser Lys Trp Lys Val Thr Ile Pro Ser Asp Leu Ala Tyr Gly
165 170 175
30
        Asp Arg Gly Ala Gly Glu His Ile Lys Pro Gly Ser Thr Leu Ile Phe
180 185 190
        Ile Ile Glu Leu Leu Ser Ile Asn Lys
35
        (2) INFORMATION FOR SEQ ID NO: 267
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 279 amino acids
40
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
45
            (iii) HYPOTHETICAL: YES
              (vi) ORIGINAL SOURCE:
                     (A) ORGANISM: Perphyromonas gingivalis
50
              (ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...279
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:267
55
        Glń Lys Asn Lys Arg Lys Met Lys Lys Ala Leu Leu Ile Gly Ala Ala 1 \phantom{a} 10 \phantom{a} 15
        Leu Leu Gly Ala Val Ser Phe Ala Ser Ala Gln Ser Leu Ser Thr Ile
20 25 30
        Lys Val Gln Asn Asn Ser Val Gln Gln Pro Arg Glu Glu Ala Thr Ile
35 40
60
        Gln Val Cys Gly Glu Leu Ala Glu Gln Val Asp Cys Ile Gly Thr Gly
50 60
        Asn Ser Ala Ile Ile Ala Ala Ala Ala Lys Phe Glu Ser Asp Asp Leu
65 70 75 80
65
        Glu Ser Tyr Val Gly Trp Glu Ile Met Ser Val Asp Phe Pro Gly
85 90
        Tyr Lys Ala Cys Lys Tyr Thr Ser Ala Val Trp Ala Asp Asp Met Thr
100 105 110
70
        Ile Leu Gly Gln Ser Glu Asp Ser Asp Pro Glu Met Gln Thr Ile Asn
115 120 125
        Asn Leu Ala Leu Lys Thr Ser Val Lys Ile Glu Ala Gly Lys Asn Tyr
130 135 140
Ile Val Gly Tyr Ile Ala Asn Thr Ala Gly Gly His Pro Ile Gly Cys
145 150 160
75
```

```
Asp Gln Gly Pro Ala Val Asp Gly Tyr Gly Asp Leu Val Ser Ile Ser 165 $170$
      Glu Asp Gly Gly Ala Thr Phe Pro Pro Phe Glu Ser Leu His Gln Ala
                  180
                                       185
                                                             190
      Val Pro Thr Leu Asn Tyr Asn Ile Tyr Val Val His Leu Lys Lys
195 200 205
      Gly Glu Gly Val Glu Ala Val Leu Thr Asn Asp Lys Ala Asn Ala Tyr
210 215 220
      Val Gln Asn Gly Val Ile Tyr Val Ala Gl; Ala Asn Gly Arg Gln Val
10
                         230
                                                 235
      Ser Leu Phe Asp Het Asn Gly Lys Val Val Tyr Thr Gly Val Ser Glu 245 255
      Thr Ile Ala Ala Pro Gln Lys Gly Met Tyr Ile Leu Arg Val Gly Ala
                  260
15
      Lys Ser Ile Lys Leu Ala Ile
       (2) INFORMATION FOR SEQ ID NO: 268
20
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 157 amino acids
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
25
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
30
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...157
35
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:268
      Arg Unk Phe Leu Pro Glu Lys Ala Leu Tyr Ile Gly Cys Arg Val Glu
1 5 10 15
40
      Thr Gln Glu Gly His Ala Val Gly Phe Gly Leu Asp Asp Gly Pro Ala 20 25 30
      Net Lys Gly Lys Gly Asp Leu Val Gly Ser Tyr Leu Pro Gly Ala Ala 35 40 45
       Pro Met Pro Phe Val Pro Leu Ser Asp Ile Pro Ala Arg Ser Met Asp
50 55 60
45
      Ala Asn Phe Tyr Ile Tyr Ser Arg Ile Ser Leu Gly Ser Gly Thr Gln
65 70 75 80
      Asp Val Leu Gln His Arg Met Lys Val Tyr Pro Asn Pro Ala Thr Thr
85 90 95
      Glu Leu His Val Glu Ala Leu Ser Ala Trp Val Gly Glu Gin Ala Ala
100 105 110
50
       Val Tyr Asp Het Arg Gly Arg Arg Val Ser Ala Arg Thr Val Asp Ser
115 120 125
      Glu Lys Leu Cys Ile Asp Ile Ala Ser Leu Pro Val Gly Val Tyr Met
130 135
55
       Leu Arg Ile Gly Ser Tyr Ser Ala Lys Phe Glu Lys Arg
       (2) INFORMATION FOR SEQ ID NO:269
60
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 562 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
65
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
70
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
75
                  (B) LOCATION 1...562
```

| | | (xi) SEQUENCE DESCRIPTION: SEQ ID NO:269 | | | | | | | | | | | | | | |
|----|------------|--|------------|-----|--------------------|------------|-----------|------------|-----|-------------|------------|-----------|-----------|-----|------------|-----|
| 5 | Thr 1 | Arg | Asn | Vāl | Ser 5 | Leu | Ile | Lys | Met | Pro 10 | Arg | Ile | Met | Lys | Leu 15 | Lys |
| J | Ile | | | 20 | | | | | 25 | Phe | | | | 30 | | |
| | Pro | Leu | Ala 35 | Lys | Ala | Gln | Met | Asp 40 | lle | Gly | Gly | Asp | Asp 45 | Val | Leu | Ile |
| 10 | Glu | Thr 50 | | Ser | Thr | Leu | Ser 55 | | Tyr | Ser | Glu | Asp 60 | | Tyr | Tyr | Γλε |
| | 65 | | | | _ | 70 | _ | - | | Туг | 75 | | | | | 80 |
| 15 | | | | | 85 | | | | | Туг 90 | | | | | 95 | |
| | | | | 100 | | | | | 105 | Gly | | | | 110 | | |
| 00 | | | 115 | | | | | 120 | | Val | | | 125 | | | |
| 20 | | 130 | | | | | 135 | | | Ala | | 140 | | | | |
| | Thr 145 | Ile | Gly | Asn | Gl y | Val 150 | Leu | Leu | Het | His | Arg 155 | His | Asp | Ala | Asp | 160 |
| 25 | Asn | Asn | Thr | Glu | Су <i>в</i> 165 | Val | туг | Lys | rys | Asp 170 | Phe | Pro | Asn | Asn | Arg 175 | Leu |
| | | - | | 180 | | | | | 185 | Arg | | | | 190 | | |
| | | | 195 | | | | | 200 | | Val | | | 205 | | | |
| 30 | | 210 | | | | | 215 | | | Phe | | 220 | | | | |
| | 225 | | | | | 230 | | | | Arg | 235 | | | | | 240 |
| 35 | | | | | 245 | | | | | Thr 250 | | | | | 255 | |
| | | | - | 260 | | | - | | 265 | Phe | | | | 270 | | |
| | Asp | Gly | Phe 275 | Asp | Ile | Gly | Phe | 11e 280 | | Asn | Phe | Val | 285 | Tyr | Asp | Pro |
| 40 | - | 290 | | - | | | 295 | | | Ile | | 300 | | | | |
| | 305 | | | | | 310 | | | | Ser | 315 | | | | | 320 |
| 45 | | | - | | 325 | | | | | Gl y 330 | | | | | 335 | |
| | | | | 340 | _ | | _ | | 345 | Val | | | | 350 | | |
| 50 | | | 355 | | | | | 360 | | rys - | | | 365 | | | |
| 50 | | 370 | | | | | 375 | | | Tyr | | 380 | | | | |
| | 385 | | | | | 390 | | | | Leu | 395 | | | | | 400 |
| 55 | | | | | 405 | | | | | Gly 410 | | | | | 415 | |
| | | | | 420 | | | | | 425 | Arg | | | | 430 | | |
| co | | | 435 | | | | | 440 | | Trp | | | 445 | | | |
| 60 | | 450 | | | | | 455 | | | Pro | | 460 | | | | |
| | 465 | | | - | | 470 | | | | Val | 475 | | | | | 480 |
| 65 | | | | | 485 | | | | | Lys 490 | | | | | 495 | |
| | _ | | _ | 500 | | | | | 505 | | | | | 510 | | |
| 70 | | | 515 | | | | | 520 | | Val | | | 525 | | | |
| 70 | - | 530 | | - | - | | 535 | | | Tyr | | 540 | - | - | | _ |
| | 545 | | | val | val | 550 | АЅР | Inr | GIU | Туг | 555 | val | ъtи | ьÀs | тте | 560 |
| 75 | val | Glu | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |

```
(2) INFORMATION FOR SEQ ID NO:270
             (i) SEQUENCE CHARACTERISTICS:
  5
                   (A) LENGTH: 391 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
 10
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
 15
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...3\overline{9}1
 20
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:270
       Gln Net Lys Arg Leu Leu Pro Phe Leu Leu Leu Ala Gly Leu Val Ala
1 5 10
                                             10
       Val Gly Asn Val Ser Ala Gln Ser Pro Arg Ile Pro Gln Val Asp Val
25
                    20
                                        25
       His Thr Arg Ile Ala Arg Asn Ala Arg Tyr Arg Leu Asp Lys Ile Ser
35 40 45
       Val Pro Asp Ser Arg Gln Ile Phe Asp Tyr Phe Tyr Lys Glu Glu Thr
50 55 60
30
       Ile Pro Thr Lys Ile Gln Thr Thr Thr Gly Gly Ala Ile Thr Ser Ile 65 70 75 80
       Asp Ser Leu Phe Tyr Glu Asp Asp Arg Leu Val Gln Val Arg Tyr Phe 85 90 95
       Asp Asn Asn Leu Glu Leu Lys Gln Ala Glu Lys Tyr Val Tyr Asp Gly
35
       Ser Lys Leu Val Leu Arg Glu Ile Arg Lys Ser Pro Thr Asp Glu Thr
       Pro Ile Lys Lys Val Ser Tyr His Tyr Leu Cys Gly Ser Asp Met Pro
130 135 140
       Phe Glu Ile Thr Thr Glu Met Ser Asp Gly Tyr Phe Glu Ser His Thr
145 150 155 160
40
       Leu Asn Tyr Leu Asn Gly Lys Ile Ala Arg Ile Asp Ile Met Thr Gln 165 170 175
       Gln Asn Pro Ser Ala Glu Leu Ile Glu Thr Gly Arg Met Val Tyr Glu
180 185 190
45
       Phe Asp Ala Asn Asn Asp Ala Val Leu Leu Arg Asp Ser Val Phe Leu
195 200 205
       Pro Leu Gln Asn Lys Trp Val Glu Met Phe Thr His Arg Tyr Thr Tyr 210 215 220
       Asp Asn Lys His Asn Cys Ile Arg Trp Glu Gln Asp Glu Phe Gly Thr
225 230 235 240
50
       Leu Thr Leu Ala Asn Asn Phe Glu Tyr Asp Thr Thr Ile Pro Leu Ser
245 250 255
       Ser Val Leu Phe Pro Thr His Glu Glu Phe Phe Arg Pro Leu Leu Pro 260 265 270
55
       Asn Phe Met Lys His Met Arg Thr Lys Gln Thr Tyr Phe Asn Asn Ser
275 280 285
       Gly Glu Gly Leu Ser Glu Val Cys Asp Tyr Asn Tyr Phe Tyr Thr Asp
290 295 300
60
      Het Gln Gly Asn Ala Leu Thr Asp Val Ala Val Asn Glu Ser Ile Lys
305 310 315 320
      Ile Tyr Pro Arg Pro Ala Thr Asp Phe Leu Arg Ile Glu Gly Ser Gln
325
330
335
      Leu Leu Arg Leu Ser Leu Phe Asp Met Asn Gly Lys Leu Ile Arg Ala
340 345 350
65
      Thr Glu Leu Thr Gly Asp Leu Ala Ile Ile Gly Val Ala Ser Leu Pro
355 360 365
      Arg Gly Thr Tyr Ile Ala Glu Ile Thr Ala Ala Asn Ser Lys Thr Ile
        370
                              375
70
       Arg Ala Lys Val Ser Leu Arg
       (2) INFORMATION FOR SEQ ID NO:271
75
            (i) SEQUENCE CHARACTERISTICS:
```

```
(A) LENGTH: 428 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
 5
           (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
10
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...428
15
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:271
       Lys Arg Asn Pro Leu Pro Leu Thr Ala Ser Asn Arg Lys Ile Phe Ile
                                                10
       Lys Het Arg Gln His Leu Ser Leu Phe Pro Phe Ile Leu Phe Leu Leu 20 25 30
20
       Leu Ala Phe Ser Tyr Val Gly Cys Arg Thr Val Arg Gln Thr Pro L;s
35 40
       Gln Ser Glu Arg Tyr Val Val Val Leu Ser Leu Asp Gly Phe Arg Pro
50 55 60
25
       Asp Tyr Thr Asp Arg Ala Arg Thr Pro Ala Leu Asp Arg Met Ala Gln 65 70 75
       Glu Gly Leu Ser Gly Ser Leu Gln Pro Cys Phe Pro Ser Leu Thr Phe
85 90 95
       Pro Asn His Tyr Ser Met Ala Thr Gly Leu Tyr Pro Asp His His Gly 100 105 110
30
       Ile Val Ala Asn Glu Phe Val Asp Ser Leu Leu Gly Ile Phe Arg Ile
115 120 125
       Ser Asp Arg Lys Ala Val Glu Thr Pro Gly Phe Trp Gly Gly Glu Pro
130 135 140
35
       Val Trp Asn Thr Ala Ala Arg Gln Gly Ile Arg Thr Gly Val Tyr Phe
145 150 155 160
       Trp Val Gly Ser Glu Thr Ala Val Asn Gly Asn Arg Pro Trp Arg Trp
165 170 175
40
       Lys Lys Phe Ser Ser Thr Val Pro Phe Arg Asp Arg Ala Asp Ser Val
       Ile Ala Trp Leu Gly Leu Pro Glu Lys Glu Arg Pro Arg Leu Leu Met
195 200 205
       Trp Tyr Ile Glu Glu Pro Asp Met Ile Gly His Ser Gln Thr Pro Glu
210 215 220
45
       Ser Pro Leu Thr Leu Ala Met Val Glu Arg Leu Asp Ser Val Val Gly
225 230 235 240
       Tyr Phe Arg Lys Arg Leu Asp Ser Leu Pro Ile Ala Ala Gln Thr Asp
245 250 255
       Phe Ile Ile Val Ser Asp His Gly Met Ala Thr Tyr Glu Asn Glu Lys
260 265 270
50
       Cys Val Asn Leu Ser His Tyr Leu Pro Ala Asp Ser Phe Leu Tyr Het
275 280 285
       Ala Thr Gly Ala Phe Thr His Leu Tyr Pro Lys Pro Ser Tyr Thr Glu
290 295 300
55
       Arg Ala Tyr Glu Ile Leu Arg Ala Ile Pro His Ile Ser Val Tyr Arg
305 310 315
       Lys Gly Glu Val Pro Lys Arg Leu Arg Cys Gly Thr Asn Pro Arg Leu
325 330 335
       Gly Glu Leu Val Val Ile Pro Asp Ile Gly Ser Thr Val Phe Phe Ala
340 345 350
60
       Ile Asn Glu Asp Val Arg Pro Gly Ala Ala His Gly Tyr Asp Asn Gln 355 360 365
       Ala Pro Glu Het Arg Ala Leu Leu Arg Ala Val Gly Pro Asp Phe Arg 370 375 380
65
        Pro Gly Ser Arg Val Glu Asn Leu Pro Asn Ile Thr 11e Tyr Pro Leu
385 390 395 400
       lle Cys Arg Leu Leu Gly Ile Glu Pro Ala Pro Asn Asp Ala Asp Glu
405 410 415
70
        Thr Leu Leu Asn Gly Leu Ile Arg Asp Lys Arg Pro
        (2) INFORMATION FOR SEQ ID NO:272
75
             (i) SEQUENCE CHARACTERISTICS:
```

```
(A) LENGTH: 282 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
  5
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
 10
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...282
 15
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:272
       Leu Ser Arg Gly Val Phe Pro Leu Met Tyr Gly Arg Arg Gly Ser Ile
20
       Arg Ala Ser Ser Gly His Arg Asp Lys Ile Phe Lys Asn Thr Ile Ile 20 25 30
       Arg Phe Ile Thr Het Lys Val Gly Leu Phe Ile Pro Cys Tyr Val Asn
35 40
       Ala Val Tyr Pro Glu Val Gly Ile Ala Thr Tyr Lys Leu Leu Lys Ser 50 60
25
       Leu Asp Ile Asp Val Asp Tyr Pro Het Asp Gln Thr Cys Cys Gly Gln
65 70 75
       Pro Het Ala Asn Ala Gly Phe Glu Gln Lys Ala Gln Lys Leu Ala Leu
85 90 95
30
       Arg Phe Glu Glu Leu Phe Glu Ser Tyr Asp Val Val Gly Pro Ser
       Ala Ser Cys Val Ala Phe Val Lys Glu Asn Tyr Asp His Ile Leu Arg
       Pro Thr Gly His Val Cys Lys Ser Ala Ala Lys Val Arg Asp Ile Cys 130 135 140
35
       Glu Phe Leu His Asp Asp Leu Lys Ile Thr Ser Leu Pro Ser Arg Phe 145 150 155 160
       Ala His Lys Val Ser Leu His Asn Ser Cys His Gly Val Arg Glu Leu
165 170 175
40
       His Leu Ser Thr Pro Ser Glu Val His Arg Pro Tyr His Asn Lys Val
                                         185
       Arg Arg Leu Leu Glu Met Val Gln Gly Ile Glu Val Phe Glu Pro Lys
               195
                                    200
       Arg Ile Asp Glu Cys Cys Gly Phe Gly Gly Het Tyr Ser Val Glu Glu 210 215 220
45
       Pro Glu Val Ser Thr Cys Met Gly His Asp Lys Val Leu Asp His Ile
225 230 235 240
       Ser Thr Gly Ala Glu Tyr Ile Thr Gly Pro Asp Ser Ser Cys Leu Met 245 250 255
50
       His Met Gln Gly Val Ile Asp Arg Glu Lys Leu Pro Ile Lys Thr Ile
                   260
                                       265
       His Ala Val Glu Ile Leu Ala Ala Asn Leu
55
       (2) INFORMATION FOR SEQ ID NO:273
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 251 amino acids
                  (B) TYPE: amino acid
60
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
65
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
70
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...251
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:273
75
      Pro Leu Lys Lys Arg Met Asp Ile Val Ser Met Ala Asp Lys Ala Leu
```

```
10
       Val Val Glu Het Arg Asp Val Thr Leu Cys Gln Glu Glu Asn Val Ile
20 25 30
       Phe Gln Asn Leu Asn Leu Thr Leu Ser Ala Gly Asp Phe Val Tyr Leu 35 40 45
 5
       Ile Gly Ser Val Gly Ser Gly Lys Ser Thr Leu Leu Lys Ala Leu Tyr 50 60
       Ala Glu Val Pro Ile Ser Ala Gly Tyr Ala Arg Val Ile Asp Tyr Asp 65 70 75 80
       Leu Ala Lys Leu Lys Arg Lys Gln Leu Pro Tyr Leu Arg Arg Asn Leu 85 90\, 95
10
       Gly Ile Val Phe Gln Asp Phe Gln Leu Leu Asn Gly Arg Thr Val Ala
100 105 110
       Glu Asn Leu Asp Phe Val Leu Arg Ala Thr Asp Trp Lys Asn Arg Ala 115 120 125
15
       Asp Arg Glu Gln Arg Ile Glu Glu Val Leu Thr Arg Val Gly Het Ser
130 135 140
       Arg Lys Ala Tyr Lys Arg Fro His Glu Leu Ser Gly Gly Glu Gln Gln 145 150 150 160
20
       Arg Val Gly Ile Ala Arg Ala Leu Leu Ala Lys Pro Ala Leu Ile Leu
165 170 175
       Ala Asp Glu Pro Thr Gly Asn Leu Asp Ser Val Thr Gly Leu Gln Ile
180 185 190
       Ala Ser Leu Leu Tyr Glu Ile Ser Lys Gl<br/>n Gly Thr Ala Val Leu Met 195$200$ 205
25
       Ser Thr His Asn Ser Ser Leu Leu Ser His Leu Pro Ala Arg Thr Leu 210 215 220
       Ala Val Arg Lys Asn Gly Asp Ala Ser Ser Leu Val Glu Leu Ser Ala
225 230 235 240
30
       Asp Ala Val Ser Arg Lys Asn Thr Glu Ile Asp
       (2) INFORMATION FOR SEQ ID NO: 274
35
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 238 amino acids(B) TYPE: amino acid
                  (D) TOPOLOGY: linear
40
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
45
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KET: misc_feature
                  (B) LOCATION 1...238
50
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:274
       Thr Arg His Cys Pro Ala Cys Arg Ser Ser Phe His Leu Ile Lys Thr 1 5 10 15
                                                10
55
       Ser Lys Thr Het Ile Glu Ile Ser Asn Leu Thr Lys Val Phe Arg Thr
20 25 30
       Glu Glu Ile Glu Thr Val Ala Leu Asp Gly Val Ser Leu Lys Val Asp
35 40 45
       Lys Gly Glu Phe Ile Ala Ile Met Gly Pro Ser Gly Cys Gly Lys Ser 50 60
60
       Thr Leu Leu Asn Ile Leu Gly Leu Leu Asp Asn Pro Thr Ser Gly Ile
65 70 80
       Tyr Lys Leu Asp Gly Ala Glu Val Gly Asn Leu Arg Glu Lys Asp Arg
85 90 95
       Thr Ala Val Arg Lys Gly Asn Ile Gly Phe Val Phe Gln Ser Phe Asn 100 105 110
       Leu Ile Glu Glu Met Thr Val Ser Glu Asn Val Glu Leu Pro Leu Val
115 120 125
       Tyr Leu Gly Val Lys Ala Ser Glu Arg Lys Glu Arg Val Glu Glu Ala
130 135 140
70
       Leu Arg Lys Met Ser Ile Ser His Arg Ala Gly His Phe Pro Asn Gln
145 150 155 160
       Leu Ser Gly Gly Gln Gln Gln Arg Val Ala Ile Ala Arg Ala Val Val
165 170 175
75
       Ala Asn Pro Lys Leu Ile Leu Ala Asp Glu Pro Thr Gly Asn Leu Asp
```

```
180
                                               185
        Ser Lys Asn Gly Ala Asp Val Het Glu Leu Leu Arg Gly Leu Asn Arg
195 200 205
        Glu Gly Ala Thr Ile Val Met Val Thr His Ser Glu His Asp Ala Arg
210 215 220
  5
        Ser Ala Gly Arg Ile Ile Asn Leu Phe Asp Gly Lys Ile Arg
        (2) INFORMATION FOR SEQ ID NO:275
 10
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 604 amino acids
                     (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
 15
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
 20
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAMAE/KEY: misc_feature
25
                    (B) LOCATION 1...604
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:275
        Ser Thr Glu Thr Asn Ser Lys Ser Glu Met Lys Glu Phe Phe Lys Het
30
                                                   1.0
        Phe Phe Ala Ser Ile Leu Gly Val Ile Thr Ala Gly Ile Ile Leu Phe 20 25 30
        Cys Ile Phe Leu Phe Ile Phe Phe Gly Ile Val Ala Gly Ile Ala Ser
35 40 45
        Lys Ala Thr Gly Gly Thr Ile Pro Lys Ile Glu Ala Asn Ser Ile Leu 50 55 60

His Ile Unk Asn Ser Ser Phe Pro Glu Ile Val Ser Ala Asn Pro Trp 65 70 75 80
35
        Ser Met Leu Thr Gly Lys Asp Glu Ser Val Ser Leu Ser Gln Ala Val
40
        Glu Ala Ile Gly Gln Ala Lys Asn Asn Pro Asn Ile Thr Gly Ile Phe
100 105 110
        Leu Asp Leu Asp Asn Leu Ser Val Gly Met Ala Ser Ala Glu Glu Leu
115 120 125
45
        Arg Arg Ala Leu Gln Asp Phe Lys Met Ser Gly Lys Phe Val Val Ser
130 135 140
        Tyr Ala Asp Arg Tyr Thr Gln Lys Gly Tyr Tyr Leu Ser Ser Ile Ala
145 150 155 160
       Asp Lys Leu Tyr Leu Asn Pro Lys Gly Met Leu Gly Leu Ile Gly Ile
165 170 175
50
       Ala Thr Gln Thr Met Phe Tyr Lys Asp Ala Leu Asp Lys Phe Gly Val
       Lys Met Glu Ile Phe Lys Val Gly Thr Tyr Lys Ala Ala Val Glu Pro
195 200 205
55
       Phe Met Leu Asn Arg Met Ser Asp Ala Asn Arg Glu Gln Ile Thr Thr 210 215 220
       Tyr Ile Asn Gly Leu Trp Asp Lys Ile Thr Ser Asp Ile Ala Glu Ser 225 230 240
Arg Lys Thr Ala Het Asp Ser Val Lys Met Phe Ala Asp Lys Gly Glu 255
60
       Het Phe Gly Leu Ala Glu Lys Ala Val Glu Het Lys Leu Val Asp Glu
260 265 270
       Leu Ala Tyr Arg Thr Asp Val Glu Lys Glu Leu Lys Lys Met Ser Gln 285
65
       Arg Gly Glu Lys Asp Glu Leu Arg Phe Val Ser Leu Ser Gln Val Leu 290 295 300
       Ala Asn Gly Pro Met Asn Lys Thr Lys Gly Ser Arg Ile Ala Val Leu
305 310 315 320
       Phe Ala Glu Gly Glu Ile Thr Glu Glu Ile Ile Lys Lys Pro Phe Asp
325 330 335
70
       Thr Asp Gly Ser Ser Ile Thr Gln Glu Leu Ala Lys Glu Ile Lys Ala
340 345 350
       Ala Ala Asp Asp Asp Ile Lys Ala Val Val Leu Arg Val Asn Ser
355 360 365
75
       Pro Gly Gly Ser Ala Phe Thr Ser Glu Gln Ile Trp Lys Gln Val Ala
```

```
370
                                  375
                                                        380
       Asp Leu Lys Ala Lys Lys Pro lle Val Val Ser Met Gly Asp Val Ala
385 390 395 400
       Ala Ser Gly Gly Tyr Tyr Ile Ala Cys Ala Ala Asn Ser Ile Val Ala
405 410 415
 5
       Glu His Thr Thr Leu Thr Gly Ser Ile Gly Ile Phe Gly Met Phe Pro 420 425 430
       Asn Phe Ala Gly Val Ala Lys Lys Ile Gly Val Asn Met Asp Val Val 435 440 445
       Gln Thr Ser Lys Tyr Ala Asp Leu Gly Asn Thr Phe Ala Pro Met Thr
450 455 460
10
       Val Glu Asp Arg Ala Leu Ile Gln Arg Tyr Ile Glu Gln Gly Tyr Asp
465 470 475 480
       Leu Phe Leu Thr Arg Val Ser Glu Gly Arg Asn Arg Thr Lys Ala Gln
485 490 495
15
       Ile Asp Ser Ile Ala Gln Gly Arg Val Trp Leu Gly Asp Lys Ala Leu 500 505 510
       Ala Leu Gly Leu Val Asp Glu Leu Gly Gly Leu Asp Thr Ala Ile Lys 515 520 525
       Arg Ala Ala Lys Leu Ala Gln Leu Gly Gly Asn Tyr Ser Ile Glu Tyr 530 540
20
       Gly Lys Thr Lys Arg Asn Phe Phe Glu Glu Leu Leu Ser Ser Ser Ala
545 550 560
       Ala Asp Met Lys Ser Ala Ile Leu Ser Thr Ile Leu Ser Asp Pro Glu
565 570 575
25
       Ile Glu Val Leu Arg Glu Leu Arg Ser Het Pro Pro Arg Pro Ser Gly
                  580
                                585
       Ile Gln Ala Arg Leu Pro Tyr Tyr Phe Met Pro Tyr
30
       (2) INFORMATION FOR SEQ ID NO: 276
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 324 amino acids
35
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
40
          (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
45
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...324
           (%i) SEQUENCE DESCRIPTION: SEQ ID NO:276
50
       Leu Thr Leu Phe Cys Cys Asn Tyr Phe Lys Gln Met Arg Ala Asn Ile
1 5 10
       Trp Gln Ile Leu Ser Val Ser Val Leu Phe Phe Gly Thr Ala Ile
                   20
                                        25
55
       Gly Gln Ala Gln Ser Arg Asn Arg Thr Tyr Glu Ala Tyr Val Lys Gln 35 40 45
       Tyr Ala Asp Glu Ala Ile Arg Gln Het Ser Arg Tyr Asn Ile Pro Ala
50 55 60
       Ser Ile Thr Ile Ala Gln Ala Leu Val Glu Thr Gly Ala Gly Ala Ser 65 70 75 80
60
       Thr Leu Ala Ser Val His Asn Asn His Phe Gly Ile Lys Cys His Lys
85 90
       Ser Trp Thr Gly Lys Arg Thr Tyr Arg Thr Asp Asp Ala Pro Asn Glu
65
       Cys Phe Arg Ser Tyr Ser Ala Ala Arg Glu Ser Tyr Glu Asp His Ser 115 120 125
       Arg Phe Leu Leu Gln Pro Arg Tyr Arg Pro Leu Phe Lys Leu Asp Arg
130 135 140
       Glu Asp Tyr Arg Gly Trp Ala Thr Gly Leu Gln Arg Cys Gly Tyr Ala
145 150 155 160
70
       Thr Asn Arg Gly Tyr Ala Asn Leu Leu Ile Lys Met Val Glu Leu Tyr
165 170 175
       Glu Leu Tyr Ala Leu Asp Arg Glu Lys Tyr Pro Ser Trp Phe His Lys
180 185 190
75
       Ser Tyr Pro Gly Ser Asn Lys Lys Ser His Gln Thr Thr Lys Gln Lys
```

```
200
        Gln Ser Gly Leu Lys His Glu Ala Tyr Phe Ser Tyr Gly Leu Leu Tyr
210 215 220
        Ile Ile Ala Lys Gln Gly Asp Thr Phe Asp Ser Leu Ala Glu Glu Phe
225 230 235 240
  5
       Asp Het Arg Ala Ser Lys Leu Ala Lys Tyr Asn Asp Ala Pro Val Asp 245 250 255
        Phe Pro Ile Glu Lys Gly Asp Val Ile Tyr Leu Glu Lys Lys His Ala 260 265 270
       Cys Ser Ile Ser Lys His Thr Gln His Val Val Arg Val Gly Asp Ser 275 280 285
 10
       Met His Ser Ile Ser Gln Arg Tyr Gly Ile Arg Met Lys Asn Leu Tyr
290 295 300
       Lys Leu Asn Asp Lys Asp Gly Glu Tyr Ile Pro Gln Glu Gly Asp Ile
305 310 315 320
 15
        (2) INFORMATION FOR SEQ ID NO:277
 20
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 533 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 25
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) HAME/KEY: misc_feature
35
                   (B) LOCATION 1...533
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:277
       Arg Ile Pro Asp Glu Gln Thr Gly Arg Ile Met Asp Gly Arg Arg Tyr
40
                                              10
       Ser Asp Gly Leu His Gln Ala Ile Glu Ala Lys Glu His Val Lys Val 20 25 30
       Glu Ala Ala Thr Gln Thr Phe Ala Thr Ile Thr Leu Gln Asn Tyr Phe 35 40 45
       Arg Met Tyr His Lys Leu Ala Gly Met Thr Gly Thr Ala Glu Thr Glu
50 55 60
45
       Ala Gl; Glu Leu Trp Asp Ile Tyr Lys Leu Asp Val Val Val Ile Pro
65 70 75
       Thr Asn Lys Pro Ile Ala Arg Lys Asp Het Asn Asp Arg Ile Tyr Lys
85 90
50
       Thr Ala Arg Glu Lys Tyr Ala Ala Val Ile Glu Glu Ile Val Arg Leu
100 105 110
       Val Leu Asn Ala Lys Leu His Gln Lys Glu Ala Glu Ile Val Ala Gln
145 150 155 160
       Ala Gly Gln Lys Gly Thr Val Thr Ile Ala Thr Asn Met Ala Gly Arg
60
       Gly Thr Asp Ile Lys Leu Ser Ala Glu Val Lys Lys Ala Gly Gly Leu
180 185 190
       Ala Ile Ile Gly Thr Glu Arg His Glu Ser Arg Arg Val Asp Arg Gln
195 200 205
65
       Leu Arg Gly Arg Ser Gly Arg Gln Gly Asp Pro Gly Ser Ser Ile Phe
210 215 220
       Tyr Val Ser Leu Glu Asp His Leu Met Arg Leu Phe Ala Thr Glu Lys
225 230 235 240
      Ile Ala Ser Leu Met Asp Arg Leu Gly Phe Lys Glu Gly Glu Val Leu
245 250 255
70
      Glu Asn Asn Het Leu Ser Lys Ser Val Glu Arg Ala Gln Lys Lys Val
260 265 270
      Glu Glu Asn Asn Phe Gly Ile Arg Lys His Leu Leu Glu Tyr Asp Asp
275 280 285
75
      Val Het Asn Ser Gln Arg Glu Val Ile Tyr Thr Arg Arg Arg His Ala
```

```
295
       Leu Met Gly Glu Arg Ile Gly Met Asp Val Leu Asn Thr Ile Tyr Asp 305 310 310
       Val Cys Lys Ala Leu Ile Asp Asn Tyr Ala Glu Ala Asn Asp Phe Glu
325 330 335
       Gly Phe Lys Glu Asp Leu Met Arg Ala Leu Ala Ile Glu Ser Pro Ile
340 345 350
       Thr Gln Glu Ile Phe Arg Gly Lys Lys Ala Glu Glu Leu Thr Asp Het
355 360 365
10
       Leu Phe Asp Glu Ala Tyr Lys Ser Phe Gln Arg Lys Met Asp Leu Ile
370 375 380
       Ala Glu Val Ala His Pro Val Val His Gln Val Phe Glu Thr Gln Ala
385 390 395 400
       Ala Val Tyr Glu Arg Ile Leu Ile Pro Ile Thr Asp Gly Lys Arg Val
405 410 415
15
       Tyr Asn Ile Gly Cys Asn Leu Arg Glu Ala Asp Glu Thr Gln Gly Lys
420 425 430
       Ser Ile Ile Lys Glu Phe Glu Lys Ala Ile Val Leu His Thr Ile Asp
435 440 445
       Glu Ser Trp Lys Glu His Leu Arg Glu Met Asp Glu Leu Arg Asn Ser
450 455 460
20
       Val Gln Asn Ala Ser Tyr Glu Asn Lys Asp Pro Leu Leu Ile Tyr Lys
465 470 475 480
       Leu Glu Ser Tyr Glu Leu Phe Arg Lys Met Val Glu Ala Met Asn Arg
485
485
490
495
Lys Thr Val Ala Ile Leu Met Arg 23
25
                               505
                                                        510
       Pro Ser Gln Glu Glu Leu Glu His Arg Arg Gln Ile Glu Ile Arg His
515 520 525
              515
30
       Ala Thr Gln Gln Arg
       (2) INFORMATION FOR SEQ ID NO:278
35
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 720 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
45
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...720
50
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:278
       Lys Ser Cys Arg Val Ile Gly Gln Thr Arg Arg Tyr Gly Cys Cys Pro
1 10 15
       Val Gly Leu Arg Pro Cys His Gln Ser Leu Ser Gly Val Cys Arg Cys
20 25 30
       Leu Phe Gln Ser Arg Pro Ala Val Ala Phe Ala Arg Lys Gly Gln Arg
35 40
       Arg His Arg Arg Ser Glu Ser Gly Arg Arg Ile Arg Ala Leu Gln Gly
50 55 60
60
       Val Gln His His Gln Thr Asn Glu His Glu Val Het Ile Ser Val Asn
65 70 75 80
       Asn Leu Thr Val Asp Phe Gly Thr Arg Leu Leu Phe Asp Gln Val Ser
85 90 95
65
       Fhe Val Ile Asn Arg Arg Asp Arg Ile Ala Leu Val Gly Lys Asn Gly
100 105 110
       Ala Gly Lys Ser Thr Leu Leu Lys Leu Ile Ala Gly Net Glu Glu Pro
115 120 125
       Thr Ser Gly His Ile Ala Arg Pro Lys Gly Ile Arg Ile Gly Tyr Leu
130 135 140
70
       Pro Gln Val Net Arg Leu Gln Asp Gly His Thr Val Tyr Glu Glu Val
145 150 155 160
       Glu Gln Ala Phe Asn Asp Ile Arg Gln Ile Glu Glu Glu Ile Arg Arg
165 170 175
75
       Leu Ser Asp Glu Met Ala Gly Arg Thr Asp Tyr Glu Ser Asp Asp Tyr
```

```
180
                                             185
        Ile Arg Leu Ile Glu His Tyr Thr Asn Met Ser Glu Thr Leu Ser Leu
195 200 205
       Met Gln Gly Asn Tyr His Ala Ala Ile Glu Gln Thr Leu Ile Gly 210 215 220
  5
       Leu Gly Phe Gly Arg Glu Asp Phe His Arg Pro Thr Ala Asp Phe Ser 225 230 235 240
       Gly Gly Trp Arg Met Arg Ile Glu Leu Ala Lys Leu Leu Gln Arg
245 250 255
       Pro Glu Val Leu Leu Asp Glu Pro Thr Asn His Leu Asp Ile Glu 260 265 270
10
       Ser Ile Gly Trp Leu Glu Gln Phe Ile Ala Thr Asn Ala Gly Ala Val
275 280 285
       Ile Leu Val Ser His Asp Arg Ala Phe Ile Asp Asn Thr Thr Arg
290 295 300
15
       Thr Ile Glu Ile Glu Leu Gly His Ile Tyr Asp Tyr Lys Thr Asn Tyr 305 310 315
       Ser His Tyr Val Glu Leu Arg Glu Glu Arg Leu Arg Gln Gln Het Arg
325 330
20
       Ala Tyr Glu Asn Gln Gln Lys Het Ile Arg Asp Thr Glu Asp Phe Ile 340 345 350
       Glu Arg Phe Arg Tyr Lys Ala Thr Lys Ser Val Gln Val Gln Ser Arg
355 360 365
       Ile Lys Gln Leu Glu Lys Val Glu Arg Val Glu Ile Asp Glu Arg Asp
370 375 380
25
       Arg Ser Ala Phe His Phe Arg Phe Ile Pro Ala Gln Pro Ser Gly Ser 385 390 395
       Tyr Pro Leu Ile Val Asp Asp Leu Ala Lys Ala Tyr Gly Asp His Gln 405 410
30
       Val Phe Ser Gly Ala Thr Tyr Thr Ile Glu Arg Gly Glu Lys Val Ala
420 430
       Phe Val Gly Lys Asn Gly Ala Gly Lys Ser Thr Met Val Lys Cys Ile
435 440 445
       Met Gly Glu Leu Thr Asp Tyr Thr Gly Lys Leu Glu Leu Gly His Asn 450 455 460
35
       Val Gln Leu Gly Tyr Phe Ala Gln Asn Glu Ala Gln Glu Leu Arg Gly 475 480
       Asp Leu Thr Val Phe Asp Thr Ile Asp Arg Glu Ala Val Gly Asp Ile
485 490 495
40
       Arg Leu Arg Leu Asn Asp Leu Leu Gl_{\rm Y} Ala Phe Leu Phe Gly Gly Glu 500 510
       Ala Ser Glu Lys Lys Val Ser Val Leu Ser Gly Gly Glu Arg Ala Arg
515 520 525
       Leu Ala Ile Ile Arg Leu Leu Gln Pro Ala Asn Phe Leu Ile Leu 530 535 540
45
       Asp Glu Pro Thr Asn His Leu Asp Met Arg Ser Lys Asp Val Leu Lys 545 550 555 560
       Glu Ala Ile Lys Asn Phe Asp Gly Thr Val Ile Val Val Ser His Asp
565 570 575
50
       Arg Glu Phe Leu Asp Gly Leu Val Ser Lys Val Tyr Glu Phe Ala Asp 580 580 590
       Gly Gln Val Asn Glu His Leu Gly Gly Ile Tyr Asp Tyr Leu Arg Thr
595 600 605
       Arg Arg Met Gln Thr Leu Thr Glu Leu Glu Arg Thr Thr Thr Ile Glu
610 615 620
55
       Thr Lys Thr Thr Arg Glu Ala Ile Pro Glu Thr Glu Ala Lys Ala Asp
625 630 635 640
       Tyr Arg Arg Gln Lys Glu Val Ala Lys Gln Leu Arg Thr Leu Glu Arg
645 655
60
       Thr Val Ala Thr Cys Glu Glu Arg Ile Gly Lys Leu Glu Ser Glu Leu
660 665 670
       Gln Ala Ile Glu Met Leu Leu Gln Asp Pro Lys His Ala Thr Asp Ala
675 680 685
       Asn Leu Phe Glu Arg Tyr Ala Gly Het Lys Gln Glu Leu Glu Lys Ala
690 695 700
65
       Het Glu Asp Trp Glu Gln Ala Ser Glu Ala Leu Ser Glu Ala Gln Gly
       (2) INFORMATION FOR SEQ ID NO:279
70
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 386 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
```

75

```
(ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
 5
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
10
                  (B) LOCATION 1...386
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:279
       Tyr Leu Lys Leu Leu Ile Leu Gln Ile Ala Leu Met Asn Phe Leu Lys
15
       Lys Glu Pro Phe Lys Ile Phe Ser Met Ile Tyr Leu Leu Leu Asp Thr
                                        25
       Ile Thr Asn Arg Ala Gly Thr Glu Arg Ala Val Ile Asn Leu Ala Asn 35 \hspace{1cm} 40 \hspace{1cm} 45
20
      Asn Leu His Ala Asn Gly His Arg Val Ser Leu Val Ser Val Cys Thr 50 60
       Lys Glu Gly Glu Pro Ser Phe Gln Val Glu Lys Gly Ile Glu Val His 65 70 75 80
      His Leu Gly Ile Arg Leu Tyr Gly Asn Ala Leu Ala Arg Lys Thr Val
85 90
25
      Tyr Phe Lys Ala Tyr Arg Arg Ile Lys Ala Leu Tyr Lys Lys Arg Glu
100 105 110
       Pro Val Leu Ile Gly Thr Asn Ile Phe Ile Asn Thr Ile Leu Ser
115 120 125
30
      Gln Ile Ser Asn Arg Gly Arg Ile Phe Thr Ile Gly Cys Glu His Ile
130 135 140
      Ser Tyr Asp Ile Ala Arg Pro Ile Thr Lys Arg Ile Arg Gly Phe Leu
145 150 160
      Tyr Ser Gly Leu Asp Ala Val Val Ala Leu Thr Lys Arg Asp Gln Gln
165 170 175
35
      Ser Phe Glu Ala Ile Leu Arg Gly Arg Ser Lys Ala Tyr Val Ile Pro
180 185 190
      Asn Gln Val Ser Phe Thr Thr Val Gln Arg Asp Ala Thr Thr His Lys
195 200 205
      Gln Met Leu Ala Ile Gly Arg Leu Thr Tyr Gln Lys Gly Phe Glu Phe 210 215 220
40
      Met Ile Glu Asp Ala Ser Arg Val Leu Arg Glu Arg Pro Asp Trp Lys 225 230 230 235
      Leu Ile Ile Val Gly Asp Gly Glu Asn Glu Ser Met Leu Arg Lys Glu 245 250 255
45
      Ile Ala Ser Arg Asn Met Glu Ser Gln Ile Glu Ile His Pro Ser Thr
260 265 270
      Pro Glu Ile Arg Lys Tyr Tyr Glu Ser Ser Ala Ile Tyr Leu Net Thr 275 280 285
50
      Ser Arg Phe Glu Gly Leu Pro Met Val Leu Leu Glu Ala Glu Ala Tyr
290 295 300
      55
      Asp Phe Ala Asp Lys Leu Arg Leu Leu Het Asp Asp Glu Thr Leu Arg
340 345 350
      Lys Lys Met Gly Gln Glu Ser Glu Leu Het Val Lys Ser Tyr Ser Pro
355 360 365
60
       Ala Asn Ile Tyr Glu Cys Trp Lys Lys Leu Phe Val Glu Ile Gly Tyr
           370
      Met Asn
       385
65
       (2) INFORMATION FOR SEQ ID NO:280
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 655 amino acids
(B) TYPE: amino acid
70
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
75
```

```
(vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
  5
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...655
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:280
 10
        Thr Thr Lys Glu Asn Glu Thr Thr Thr Lys Asn Glu Tyr Arg Ile Het
                                                1.0
       Gly Lys Ile Ile Gly Ile Asp Leu Gly Thr Thr Asn Ser Cys Val Ser
20 25 30
       Val Leu Glu Gly Asn Glu Pro Ile Val Ile Thr Asn Ser Glu Gly Lys
35 40 45
15
       Arg Thr Thr Pro Ser Val Val Ala Phe Val Asp Gly Glu Arg Lys 50 60
       Val Gly Asp Pro Ala Lys Arg Gln Ala Ile Thr Asn Pro Thr Lys Thr
65 70 80
20
       Ile Tyr Ser Ile Lys Arg Phe Met Gly Glu Thr Tyr Asp Gln Val Ser
       Arg Glu Val Glu Arg Val Pro Phe Lys Val Val Arg Gly Asp Asn Asn 100 105 110
       Thr Pro Arg Val Asp Ile Asp Gly Arg Leu Tyr Thr Pro Gln Glu Ile
115 120 125
25
       Ser Ala Het Ile Leu Gln Lys Met Lys Lys Thr Ala Glu Asp Tyr Leu 130 135 140
       Gly Gln Glu Val Thr Glu Ala Val Ile Thr Val Pro Ala Tyr Phe Asn
145 150 155 160
30
       Asp Ala Gln Arg Gln Ala Thr Lys Glu Ala Gly Glu Ile Ala Gly Leu
165 170 175
       Lys Val Arg Arg Ile Val Asn Glu Pro Thr Ala Ala Ser Leu Ala Tyr
180 185 190
       Gly Leu Asp Lys Ser Asn Lys Asp Met Lys Ile Ala Val Phe Asp Leu
195 200 205
35
       40
       Phe Asp His Val Ile Ile Asp Trp Leu Ala Glu Glu Phe Lys Ser Gln 245 250 255
       Glu Gly Val Asp Leu Arg Gln Asp Pro Met Ala Met Gln Arg Leu Lys
260 265 270
       Glu Ala Ala Glu Lys Ala Lys Ile Glu Leu Ser Ser Thr Ser Ser Thr
275 280 285
45
       Glu Ile Asn Leu Pro Tyr Ile Met Pro Val Asn Gly Ile Pro Lys His
290 295 300
       Leu Val Het Thr Leu Thr Arg Ala Lys Phe Glu Gln Leu Ala Asp Arg
305 310 320
                                                 315
50
       Leu Ile Gin Ala Cys Val Ala Pro Cys Glu Thr Ala Leu Lys Asp Ala
325 330 335
       Gly Met Ser Arg Gly Asp Ile Asp Glu Val Ile Leu Val Gly Gly Ser
340 345 350
       Thr Arg Ile Pro Ala Ile Gln Glu Ile Val Glu Lys Ile Phe Gly Lys
355 360 365
55
       Ala Pro Ser Lys Gly Val Asn Pro Asp Glu Val Val Ala Val Gly Ala 370 375 380
       Ala Ile Gln Gly Gly Val Leu Thr Gly Glu Val Lys Asp Val Leu Leu 385 390 395 400
60
       Leu Asp Val Thr Pro Leu Ser Leu Gly Ile Glu Thr Met Gly Gly Val
405 410 415
       Het Thr Arg Leu Ile Asp Ala Asn Thr Thr Ile Pro Thr Lys Lys Ser 420 425 430
       Glu Ile Phe Thr Thr Ala Val Asp Asn Gln Pro Ser Val Glu Ile His 435 440 445
65
       Val Leu Gln Gly Glu Arg Ser Leu Ala Lys Asp Asn Lys Ser Ile Gly 450 455 460
       Arg Phe Asn Leu Asp Gly Ile Ala Pro Ala Pro Arg Gln Thr Pro Gln 465 470 475 480
70
       lle Glu Val Thr Phe Asp Ile Asp Ala Asn Gly Ile Leu Asn Val Thr
485 490 495
       Ala His Asp Lys Ala Thr Gly Lys Lys Gln Asn Ile Arg Ile Glu Ala
500 505 510
       Ser Ser Gly Leu Ser Asp Asp Glu Ile Lys Arg Het Lys Glu Glu Ala
515 520 525
75
```

```
Gln Ala Asn Ala Glu Ala Asp Lys Lys Glu Lys Glu Arg Ile Asp Lys
                                 535
           530
      Ile Asn Gln Ala Asp Ser flet Ile Phe Gln Thr Glu Lys Gln Leu Lys 545 550 560
      Glu Leu Gly Asp Lys Phe Pro Ala Asp Lys Lys Ala Pro Ile Asp Thr
565 570 575
 5
      Ala Leu Asp Lys Leu Lys Glu Ala His Lys Ala Gln Asp Val Ala Ala
580 590
      Ile Asp Thr Ala Net Ala Glu Leu Gln Thr Ala Leu Ser Ala Ala Gly 595 600 605
10
      Glu Glu Leu Tyr Lys Asn Ala Gly Ala Ala Gln Gly Gly Ala Gln Pro
610 615 620
       Gly Pro Asp Phe Gly Gly Ala Gln Gly Pro Ser Ala Gly Asp Gln Pro
625 630 635 640
       Ser Asp Asp Lys Asn Val Thr Asp Val Asp Phe Glu Glu Val Lys
15
       (2) INFORMATION FOR SEQ ID NO:281
20
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 467 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
30
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...467
35
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:281
       Lys Trp Ala Arg Thr Thr Thr Leu Arg Ile Ser Asn Ala Thr Ile Tyr
       Lys Thr Het Arg Tyr Asp Leu Ala Ile lle Gly Gly Gly Pro Ala Gly
20 25 30
40
       Tyr Thr Ala Ala Glu Arg Ala Ala Lys Gly Gly Leu Lys Thr Leu Leu
35 40
       Ile Glu Lys Asn Ala Leu Gly Gly Val Cys Leu Asn Glu Gly Cys Ile
50 55 60
45
       Pro Thr Lys Thr Leu Leu Tyr Ser Ala Lys Val Leu His Gln Ile Ala 65 70 75 80
       Thr Ala Ser Lys Tyr Ala Val Ser Gly Thr Ala Asp Gly Leu Asp Leu
85 90 95
                        85
50
       Gly Lys Val Ile Ala Arg Lys Gly Lys Ile Ile Arg Lys Leu Thr Ala 100 $105$
       Gly Ile Arg Ser Arg Leu Thr Glu Ala Gly Val Glu Met Val Thr Ala
115 120 125
       Glu Ala Thr Val Thr Gly Cys Asp Ala Asp Gly Ile Ile Gly Ile Thr
130 135 140
55
       Ala Gly Glu Ala Gln Tyr Lys Ala Ala Asn Leu Leu Cys Thr Gly
145 150 155 160
       Ser Glu Thr Phe Ile Pro Pro Ile Pro Gly Val Glu Gln Thr Glu Tyr
165 170 175
       Trp Thr Asn Arg Glu Ala Leu Gln Asn Lys Glu Ile Pro Thr Ser Leu
180 185 190
60
       Val Ile Ile Gly Gly Gly Val Ile Gly Met Glu Phe Ala Ser Phe Phe
195 200 205
       Asn Gly Tie Gly Thr Gln Val His Val Val Glu Met Leu Pro Glu Ile
210 215 220
65
       Leu Asn Gly Ile Asp Pro Glu His Ala Ala Het Leu Arg Ala His Tyr
225 230 235 240
       Glu Lys Glu Gly 11e Lys Phe Tyr Leu Gly His Lys Val Thr Ser Val
245 250 255
       Arg Ash Gly Ala Val Thr Val Glu Tyr Glu Gly Glu Ser Lys Glu Ile
260 265 270
70
       Glu Gly Glu Arg Ile Leu Met Ser Val Gly Arg Arg Pro Val Leu Gln
275 280 285
       Gly Phe Glu Ser Leu Gly Leu Val Leu Ala Gly Lys Gly Val Lys Thr
 75
                                 295
```

```
Asn Glu Arg Met Gln Thr Ser Leu Pro Asn Val Tyr Ala Ala Gly Asp
                               310
                                                     315
        The Thr Gly Phe Ser Leu Leu Ala His Thr Ala Val Arg Glu Ala Glu
                       325
                                              330
                                                                        335
  5
        Val Ala Val Asp Gln Ile Leu Gly Lys Thr Asp Glu Thr Met Ser Tyr
340 345 350
        Arg Ala Val Pro Gly Val Val Tyr Thr Asn Pro Glu Val Ala Gly Val
355
360
365
       Gly Glu Thr Glu Glu Ser Leu Arg Lys Ala Gly Arg Ala Tyr Thr Val

370 375 380

Arg Arg Leu Pro Met Ala Phe Ser Gly Arg Phe Val Ala Glu Asn Glu

385 390 395 400
 10
       Gln Gly Asn Gly Glu C;s Lys Leu Leu Leu Asp Glu Glu Asn Arg Leu
405 415
       Ile Gly Ala His Leu Ile Gly Asn Pro Ala Gly Glu Leu Ile Val Thr
        Ala Ala Met Ala Ile Glu Thr Gly Met Thr Asp Arg Gln Ile Glu Arg
435 440 445
        Ile Ile Phe Pro His Pro Thr Val Gly Glu Ile Leu Lys Glu Thr Leu
20
          450
                                   455
        (2) INFORMATION FOR SEQ ID NO:282
25
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 945 amino acids(B) TYPE: amino acid
                   (D) TOPOLOGY: linear
30
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
35
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
40
                   (B) LOCATION 1...9\overline{4}5
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:282
       Pro Lys Ile Leu Met Glu Leu Lys Arg Phe Leu Ser Leu Gly Leu Leu
45
                                                10
       Leu Val Gly Phe Ile Pro Met Lys Leu Ser Ala Gln Gln Ala Gln Pro
                    20
                                           25
       Leu Pro Thr Asp Pro Ala Val Arg Val Gly Lys Leu Asp Asn Gly Leu 35 40 45
       Thr Tyr Phe Ile Arg His Asn Glu Asn Pro Lys Asp Arg Ala Asp Phe 50 60
50
       Phe Ile Ala Gln Lys Val Gly Ser Ile Leu Glu Glu Asp Ser Gln Ser 65 70 75 80
       Gly Leu Ala His Phe Leu Glu His Met Ala Phe Asn Gly Thr Lys Asn
85 90
55
       Phe Pro Gly Lys Asn Leu Ile Asn Tyr Leu Glu Thr Ile Gly Val Arg
       Phe Gly Gln Asn Leu Asn Ala Ser Thr Gly Phe Asp Lys Thr Glu Tyr
115 120
60
       Thr Ile Met Asp Val Pro Thr Thr Arg Gln Gly Ile Ile Asp Ser Cys
130 135 140
       Leu Leu Ile Leu His Asp Trp Ser Asn Asn Ile Thr Leu Asp Gly His
150 155 160
      Glu Ile Asp Glu Glu Arg Gly Val Ile Gln Glu Glu Trp Arg Ala Arg
165 170 175
65
      Arg Asp Ala Asn Leu Arg Met Phe Glu Ala Ile Leu Ala Lys Ala Met 180 185 190
       Pro Gly Asn Lys Tyr Ala Glu Arg Met Pro Ile Gly Leu Met Asp Val
70
       Val Leu Asn Phe Lys His Asp Glu Leu Arg Asn Tyr Tyr Lys Lys Trp
210 215 220
      Tyr Arg Pro Asp Leu Gln Gly Leu Val Ile Val Gly Asp Ile Asp Val 225 230 235
      Asp Tyr Val Glu Asn Lys Ile Lys Glu Leu Phe Lys Asp Val Pro Ala
75
```

| | Pro | Val | Asn | Pro 260 | Ala | Glu | Arg | Ile | Tyr 265 | Thr | Pro | Val | Glu | Asp 270 | Asn | Asp |
|------------|------------|------------|-------------|------------|------------|------------|------------|-------------------|------------|-------------|------------|------------|------------|------------|------------|------------|
| | Glu | Pro | Ile 275 | Val | Ala | Ile | Ala | Thr 280 | | Ala | Glu | Ala | Thr 285 | | Thr | Gln |
| 5 | | 290 | Ile | | | | 295 | Asp | | | | 300 | | | | |
| | 305 | | | | | 310 | | Asp | | | 315 | | | | | 320 |
| 10 | | | | | 325 | | | Glu | | 330 | | | | | 335 | |
| | | | | 340 | | | | Phe | 345 | | | | | 350 | | |
| 15 | | _ | 355 | | | | | Val 360 Val | | | | | 365 | | | |
| 10 | | 370 | | | | | 375 | Tyr | | | | 380 | | | | |
| | 385 | | | | | 390 | | Asn | | | 395 | | | | | 400 |
| 20 | | | | | 405 | | | Thr | | 410 | | | | | 415 | |
| | | | | 420 | | | | Gln | 425 | | | | | 430 | | |
| 25 | | | 435 | | | | Asn | 440 Gln | | | | Gln | 445 | | | |
| | | 450 Lys | Asn | Ala | Val | | 455 Thr | Leu | Thr | Gly | Pro | 460 Ser | Lys | Ala | Glu | Ala |
| 30 | 465 Lys | Ile | Pro | Ser | | 470 Ala | Asp | Phe | Leu | | 475 Ala | Phe | Lys | Ala | Ala 495 | 480 Arg |
| 30 | Gl n | Gln | Lys | Val 500 | 485 Glu | Ala | Lys | Lys | Asp 505 | 490 Glu | Val | Ser | Asp | Gln 510 | | Leu |
| | Het | Glu | Lչ։s 515 | | Pro | Lys | Ala | Gly 520 | | Ile | Val | Ser | Glu 525 | | Lys | Asp |
| 35 | | 530 | Phe | | | | 535 | Leu | | | | 540 | | | | |
| | 545 | | | | | 550 | | Lys | | | 555 | | | | | 560 |
| 40 | | | | | 565 | | | Ser | | 5 70 | | | | | 575 | |
| | | | | 580 | | | | Val | 585 | | | | | 590 | | |
| 45 | | | 595 | | | | | Thr 600 Gly | | | | | 605 | | | |
| 10 | | 610 | | | | | 615 | Leu | | | | 620 | | | | |
| | 625 | | | | | 630 | | Lys | | | 635 | | | | | 640 |
| 50 | | | | | 645 | | | Asn | | 650 | | | | | 655 | |
| | | | | 660 | | | | Asp | 665 | | | | Lys | 670 Pro | | |
| 55 | Ala | | | Val | Glu | Lys | Val | 680 Asn | Tyr | Asp | Gln | Val | 685 Met | | Phe | Tyr |
| | Asn 705 | | | Phe | Ala | | 695 Ala | Gly | Asp | Phe | Met 715 | 700 Phe | Phe | Phe | Ile | Gly 720 |
| 60 | | | Asp | Glu | Ala 725 | 710 Lys | Met | Lys | Pro | Leu 730 | | Glu | Thr | Туr | Leu 735 | |
| 00 | Ser | Leu | Pro | Asn 740 | | Lys | Arg | Gly | Asp 745 | | 11et | Asn | Lys | Ala 750 | | Val |
| | Pro | Ala | Ala 755 | Arg | Ser | Glу | Lys | Ile 760 | Asp | Cys | ras | Phe | Glu 765 | Lys | Glu | Het |
| 65 | | 770 | | | | | 775 | Phe | | | | 780 | | | | |
| | 785 | | | _ | | 790 | | Leu | | | 795 | | | | | 800 |
| 70 | | | | | 805 | | | Vāl | | 810 | | | | | 815 | |
| | | | | 820 | | | | Leu | 825 | | | | | 830 | | |
| <i>7</i> 5 | | | 835 | | _ | | | Thr 840 | | | | | 845 | | | |
| , 0 | ASU | VIG | тте | val | tue | VIq | GIU | Leu | GIU | ے رب | nen | VIG | בינים | OIU | ⊕r λ | 110 |

```
855
                                                         860
       Asn Val Glu Tyr Phe Lys Lys Thr Ile Glu Asn Leu Asn Lys Gln His
865 870 875 880
       Lys Glu Ser Leu Arg Glu Asn Arg Phe Trp Leu Glu Ala Met Lys Ala
885 890 895
       Ser Phe Phe Glu Gly Asn Asp Phe Ile Thr Asp Tyr Glu Ser Val Leu 900 905 910
       Asn Gly Leu Thr Pro Ala Glu Leu Gln Lys Phe Ala Ala Asp Leu Leu 915 920 925
10
       Lys Gln Gln Asn Arg Val Val Wet Het Ala Pro Val Ala Lys Ala
       Gln
15
       (2) INFORMATION FOR SEQ ID NO: 283
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 686 amino acids
                   (B) TYPE: amino acid
20
                   (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...686
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:283
35
       Phe Pro Val Ile Lys Lys Phe Leu Tyr Ser Asp His Glu Ile Phe Leu
20 25 30
       Arg Glu Ile Val Ser Asn Ala Val Asp Ala Thr Gln Lys Leu Lys Thr 35 40 45
40
       Leu Thr Ser Val Gly Glu Phe Lys Gly Glu Thr Gly Asp Leu Arg Val 50 60
       Thr Val Ser Val Asp Glu Val Ala Arg Thr Ile Thr Val Ser Asp Arg 65 70 75 80
       Gly Val Gly Net Thr Glu Glu Glu Val Glu Lys Tyr Ile Asn Gln Ile
85 90
45
       Ala Phe Ser Ser Ala Glu Glu Phe Leu Glu Lys Tyr Lys Asp Asp Lys
100 105 110
       Ala Ala Ile Ile Gly His Phe Gly Leu Gly Phe Tyr Ser Ala Phe Het
115 120 125
       Val Ser Glu Arg Val Asp Val Ile Thr Arg Ser Phe Arg Glu Asp Ala
130 135 140
       Thr Ala Val Lys Trp Ser Cys Asp Gly Ser Pro Glu Tyr Thr Leu Glu
145 150 150 160
55
       Pro Ala Asp Lys Ala Asp Arg Gly Thr Asp Ile Val Met His Ile Asp 165 170 175
       Glu Glu Asn Ser Glu Phe Leu Lys Lys Glu Lys Ile Glu Gly Leu Leu
180 185 190
       Gly Lys Tyr Cys Lys Phe Leu Thr Val Pro Ile Ile Phe Gly Lys Lys
195 200 205
60
       Gln Glu Trp Lys Asp Gly Lys Met Gln Asp Thr Asp Glu Asp Asn Gln
210 215 220
       Ile Asn Asp Thr His Pro Ala Trp Thr Lys Lys Pro Ala Asp Leu Lys
225 230 235 240
65
       Asp Glu Asp Tyr Lys Glu Phe Tyr Arg Ser Leu Tyr Pro Met Ser Glu
245 250 255
       Glu Pro Leu Phe Trp Ile His Leu Asn Val Asp Tyr Pro Phe Asn Leu
260 265 270
       Thr Gly Ile Leu Tyr Phe Pro Lys Ile Lys Asn Asn Leu Asp Leu Gln 275 280 285

Arg Asn Lys Ile Gln Leu Tyr Cys Asn Gln Val Tyr Val Thr Asp Glu 290 295 300
70
       Val Gln Gly Ile Val Pro Asp Phe Leu Thr Leu Leu His Gly Val Ile
305 310 315
75
       Asp Ser Pro Asp Ile Pro Leu Asn Val Ser Arg Ser Tyr Leu Gln Ser
```

```
325
      Asp Ala Asn Val Lys Lys Ile Ser Ser His Ile Thr Lys Lys Val Ala
340 345 350
      Asp Arg Leu Glu Glu Ile Phe Lys Asn Asp Arg Pro Thr Phe Glu Glu 355 360 365
 5
       Lys Trp Asp Ser Leu Lys Leu Phe Val Glu Tyr Gly Met Leu Thr Asp
370 375 380
       Glu Lys Phe Tyr Glu Arg Ala Ala Lys Phe Phe Leu Phe Thr Asp Net
385 390 395 400
      Asp Gly His Lys Tyr Thr Phe Asp Glu Tyr Arg Thr Leu Val Glu Gly 405 410 415
10
       Val Gln Thr Asp Lys Asp Gly Gln Val Val Tyr Leu Tyr Ala Thr Asp
420 430
      Lys His Gly Gln Tyr Ser His Val Lys Arg Ala Ser Asp Lys Gly Tyr
435 440 445
15
       Ser Val Met Leu Leu Asp Gly Gln Leu Asp Pro His Ile Val Ser Leu 450 455 460
       Leu Glu Gln Lys Leu Glu Lys Thr His Phe Val Arg Val Asp Ser Asp
465 470 475 480
20
      Thr Ile Asn Asn Leu Ile Arg Lys Glu Glu Arg Ala Glu Val Lys Leu
485 490 495
       Ser Asp Thr Glu Arg Ala Thr Leu Val Lys Leu Phe Glu Ala Arg Leu 500 505 510
       Pro Arg Asp Glu Lys Lys His Phe Asn Val Ala Phe Glu Ser Leu Gly 515 520 525
25
       Ala Glu Gly Glu Ala Ile Leu Ile Thr Gln Ala Glu Phe Met Arg Arg 530 535 540
       Het Arg Asp Het Ala Gln Leu Gln Pro Gly Het Ser Phe Tyr Gly Glu
545 550 555 560
       Leu Pro Asp Ser Tyr Asn Leu Val Leu Asn Thr Asp His Pro Leu Ile
565 570 575
30
       Asp Arg Val Leu Ser Gly Glu Lys Glu Ser Val Glu Pro Ser Leu Thr 580 585 590
       Glu Leu Arg Ala Lys Ile Ala Glu Leu Lys Ala Glu Glu Ala Lys Leu
595 600 605
35
       Leu Asp Glu Glu Lys Gly Lys Lys Pro Glu Glu Ile Pro Val Ala Thr
610 615 620
       Lys Glu Ala Lys Glu Asn Asn Ala Val Glu Gln Ala Lys Thr Glu Gly 625 630 635 640
40
       Ser Ile Asn Asp Gln Leu Thr Lys Tyr Ala Gln Asp Asn Glu Leu Ile
645 650 655
       Gly Gln Leu Ile Asp Leu Ala Leu Leu Gly Ser Gly Leu Leu Thr Gly 660 670
       Glu Ala Leu Ala Glu Phe Ile Arg Arg Ser Gln Arg Leu Leu
45
       (2) INFORMATION FOR SEQ ID NO:284
             (i) SEQUENCE CHARACTERISTICS:
50
                   (A) LENGTH: 482 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
55
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
60
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...482
65
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:284
       Asp Ile Arg Gln Lys Arg Pro Cys Phe Asn Ala Asn Leu Tyr Phe Tyr
                                               10
       Arg Ala Met Glu Lys Leu Ile Asp Ile Leu Val Val Asp Asp Asp Val
25 30
70
       Ala Val Cys Ala Ala Leu Arg Leu Val Leu Lys Arg Ala Gly Tyr Asn
35 40 45
        Pro Val Ile Ala Asn Ser Pro Asp Glu Ala Leu Ser Ile Net Arg Asn
                                                         60
75
        Pro Asp Gly Gly Cys Lys Pro Ala Val Ile Leu Met Asp Met Asn Phe
```

```
70
        Ser Leu Ser Thr Ser Gly Arg Glu Gly Leu Glu Leu Glu Lys Met
85 90 95
        Gln Ile Phe Thr Ser Cys Pro Val Ile Leu Het Thr Ala Trp Ala Ser
100 105 110
  5
        Ile Pro Leu Ala Val Glu Gly Met Arg Leu Gly Ala Phe Asp Phe Ile
115 120 125
        Gly Lys Pro Trp Asp Asn Asp Arg Leu Leu Arg Thr Ile Asp Thr Ala
130 135 140
 10
        Leu His Leu Ala Ala Pro Ser Ala Val Ala Asn Pro Ser Glu Gln Ser
145 150 155 160
                                                   155
       Asp Arg Asp Thr Ala Arg Gln Pro Lys Ala Thr Val Gln Glu Asn Asp
165 170 175
        Pro Cys Ala His Ile Ile Gly Arg Ser Asp Ala Ile Cys Lys Ile Lys
180 185 190
 15
       Glu Arg Ile Arg Arg Ile Ala Pro Thr His Ala Ser Val Leu Ile Thr
195 200 205
                                                              205
       Gly Glu Ser Gly Thr Gly Lys Glu Leu Ile Ala Glu Ala Leu His Arg 210 220 220 230 235 240 240
20
       Ile Pro Glu Ser Leu Phe Glu Ser Glu Leu Phe Gly His Lys Lys Gly 245 250 255
       Ala Phe Thr Asn Ala Phe Ser Asp Arg Lys Gly Arg Phe Glu Leu Ala 260 265 270
25
       Asp Gly Gly Thr Ile Phe Leu Asp Glu Ile Gly Glu Leu Pro Val Gly 275 280 285
       Asn Gln Val Lys Leu Leu Arg Val Leu Gln Glu Gln Thr Phe Glu Pro
290 295 300
30
       Leu Gly Glu Ser Val Ser His Arg Val Asp Ile Arg Val Val Ser Ala
305 310 315 320
       Thr Asn Ala Ser Leu Glu Arg Met Val Ala Glu Gly Arg Phe Arg Glu
325 330 335
       Asp Leu Tyr Tyr Arg Ile Asn Leu Ile His Leu His Leu Pro Pro Leu
340 350
35
       Arg Glu Arg Gln Glu Asp Ile Gln Leu Leu Val Glu Ala Phe Ser Glu
355 360 365
       Ala Phe Ala Gln Ser Asn Gly Leu Pro His Ala Val Trp Ser Ala Glu
370 375 380
                                                      380
40
       Ala Het Arg Arg Ile Cys Ala Met Pro Leu Pro Gly Asn Val Arg Glu
385 390 395 400
       Leu Lys Asn Val Val Glu Arg Thr Leu Leu Leu Ser Gly Ser Arg Glu
405 410 415
       Ile Ser Ala Arg Asp Val Ala Asp Phe Gly Ser Gln Val Thr Ala Ala 420 425 430
45
       Asp His Ser Asp Glu Arg Ala Leu Thr Asp Het Glu Glu Ala Ala Ile
435 440 445
       Arg Glu Thr Leu Thr Lys Tyr Asn Gly Asn Val Ser Arg Ala Ala Arg
450 455 460
50
       Ala Leu Gly Leu Ser Arg Ala Ala Leu Tyr Arg Arg Met Glu Lys Tyr
55
       (2) INFORMATION FOR SEQ ID NO: 285
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 263 amino acids
                   (B) TYPE: amino acid
60
                   (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
65
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
70
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...263
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:285
75
      Arg Ser Leu Gln Ser Phe Gln Asn Lys Lys His Ser Ser Met Leu Lys
```

```
10
      The Lys Asn Leu His Ala Thr Val Gln Gly Lys Glu Ile Leu Lys Gly 20 25 30
      Ile Asn Leu Glu Ile Asn Ala Gly Glu Ile His Ala Ile Met Gly Pro 35 40 45
      Asn Gly Ser Gly Lys Ser Thr Leu Ser Ser Val Leu Val Gly His Pro 50 55 60
      Ser Phe Glu Val Thr Glu Gly Glu Val Thr Phe Asn Gly Ile Asp Leu 65 70 75 80
       Leu Glu Leu Glu Pro Glu Glu Arg Ala His Leu Gly Leu Phe Leu Ser
85 90 95
10
       Phe Gln Tyr Pro Val Glu Ile Pro Gly Val Ser Met Val Asn Phe Met 100 105 110
       Arg Ala Ala Val Asn Glu His Arg Lys Ala Ile Gly Ala Glu Pro Val
115 120 125
15
       Ser Ala Ser Asp Phe Leu Lys Met Met Arg Glu Lys Arg Ala Ile Val
130 135 140
       Glu Leu Asp Asn Lys Leu Ala Ser Arg Ser Val Asn Glu Gly Phe Ser
145 150 155 160
      Gly Gly Glu Lys Lys Arg Asn Glu Ile Phe Gln Met Ala Met Leu Glu
165 170 175
20
       Pro Lys Leu Ala Ile Leu Asp Glu Thr Asp Ser Gly Leu Asp Ile Asp
180 185 190
       Ala Leu Arg Ile Val Ala Gly Gly Val Asn Arg Leu Arg Ser Pro Glu
195 200 205
25
       Asn Ala Ala Ile Val Ile Thr His Tyr Gln Arg Leu Leu Glu Tyr Ile
210 215 220
       Lys Pro Asp Phe Val His Val Leu Tyr Lys Gly Arg Ile Val Lys Ser
225 230 235 240
30
       Gly Gly Ala Glu Leu Ala Leu Thr Leu Glu Glu Lys Gly Tyr Asp Trp
       Ile Lys Glu Glu Ile Gly Glu
                    260
35
       (2) INFORMATION FOR SEQ ID NO:286
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 462 amino acids
(B) TYPE: amino acid
40
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
45
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
50
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...462
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:286
55
       Ser Met Ala Lys Glu Lys Thr Ile Tyr Val Cys Arg Ser Cys Gly Thr
       Lys Tyr Ala Lys Trp Gln Gly Asn Cys Asn Ala Cys Gly Glu Trp Asn
20 25 30
       Cys Ile Asp Glu Glu Lys Val Pro Ala Pro Ala Ser Gly Lys His Ala 35 \phantom{\bigg|}40\phantom{\bigg|}
60
       Ala Lys Ser Phe Met Pro Arg Glu Gln Asp Asn Arg Pro Arg Leu Leu
50 55 60
       Gln Asp Val Glu Ser Gly Asp Glu Glu Arg Ile Arg Leu Gly Asp Glu
65 70 75 80
       Glu Phe Asp Arg Val Leu Gly Gly Gly Ile Val Lys Gly Ala Phe Val 85 90 95
65
       Leu Leu Gly Gly Glu Pro Gly Ile Gly Lys Ser Thr Leu Ile Leu Gln 100 105 110
       Thr Val Leu Arg Leu Pro Gln Leu Arg Thr Leu Tyr Val Ser Gly Glu
70
       Glu Ser Ala Arg Gln Leu Lys Met Arg Ala Glu Arg Leu Gly Gln Ala
130 135 140
       Het Asn Gly Cys Tyr Val Tyr Cys Glu Thr Asn Ile Glu Arg Ile Leu
145 150 160
75
       Ser Arg Ala Glu Glu Leu Thr Pro Asp Leu Leu Val Ile Asp Ser Ile
```

```
Gln Thr Val Tyr Thr Glu Glu Met Glu Ser Ser Ala Gly Ser Val Gly 180 185 190
       Gln Ile Arg Glu Cys Ala Ala Leu Leu Leu Lys Tyr Cys Lys Thr Thr
195 200 205
       Gly Ile Pro Val Ile Val Ile Gly His Ile Thr Lys Glu Gly Ser Ile
210 215 220
       Ala Gly Pro Lys Val Leu Glu His Ile Val Asp Thr Val Leu Leu Phe 225 230 235 240
10
       Asp Gly Asp Lys His His Leu Tyr Arg Ile Leu Arg Gly Gln Lys Asn
245 250 255
       Arg Tyr Gly Ser Thr Ser Glu Leu Gly Ile Tyr Glu Met Arg Gln Asp
260 265 270
       Gly Leu Arg Gly Val Glu Asn Pro Ser Glu His Leu Ile Thr Arg Asn
275 280 285
15
       Arg Glu Asp Leu Ser Gly Ile Ala Ile Ala Val Ala Het Glu Gly Ile
290 295 300
       Arg Pro Ile Leu Ile Glu Ala Gln Ala Leu Val Ser Ser Ala Ile Tyr 305 310 315 320
Ala Asn Pro Gln Arg Ser Ala Thr Gly Phe Asp Ile Arg Arg Het Asn 325 330 335
20
       Met Leu Leu Ala Val Leu Glu Lys Arg Ala Gly Phe Lys Leu Ile Gln 340 345 350
       Lys Asp Val Phe Leu Asn Ile Ala Gly Gly Ile Lys Ile Ala Asp Pro 355 \hspace{1cm} 360 \hspace{1cm} 365
25
       Ala Thr Asp Leu Ala Val Ile Ser Ala Val Leu Ala Ser Ser Leu Asp
370 375 380
       Ile Val Ile Pro Pro Ala Val Cys Met Thr Gly Glu Val Gly Leu Ser 385 390 395 400
30
       Gly Glu Ile Arg Pro Val Ser Arg Ile Glu Gln Arg Ile Thr Glu Ala 405 410 415
       Arg Arg Ile Gly Phe Lys Glu Ile Leu Val Pro Ala Asp Asn Phe Arg
420 425 430
       Gln Glu Asp Ala Gly Arg Phe Gly Ile Arg Leu Val Pro Val Arg Lys
435 440 445
35
       Val Glu Glu Ala Phe Arg His Leu Phe Ser Lys Gly Arg Glu
450 455 460
       (2) INFORHATION FOR SEQ ID NO:287
40
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 373 amino acids(B) TYPE: amino acid
                    (D) TOPOLOGY: linear
45
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
50
            (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                    (A) NAME/KEY: misc feature
55
                    (B) LOCATION 1...373
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:287
       Gly Ser Cys Gly Ser Ser Pro Ala Ile Gly Arg Ser Gly Cys Ser Arg 1 \phantom{000} 5 \phantom{000} 10 \phantom{000} 15
60
       Cys Asp Arg Arg Ile Gly His Gly Leu Ser Thr Ser Ala Phe Arg Thr 20 25 30
       Tyr Ala Gly Leu Val Cys His Ser Arg Leu Arg Asp Gln Ser Ser Cys
35 40 45
       Arg Val Cys Gly Net Arg Pro Ile Gly Gln Leu Phe Pro Ser Cys
50 55
65
       Pro Arg Arg Ser Ala Ser Asp Ala Gly Arg Thr Glu Arg Ile Gln Ser 65 70 75 80
       Pro Leu Gln Asn Leu Leu Gln Glu Met Gln Phe Ser Pro Thr Asn Arg
70
                         85
                                              90
       Arg Asn Pro Phe Asp Asn Met Asn Ser Arg His Leu Thr Ile 100 105 110
       Ile Ala Gly Leu Ser Leu Phe Val Leu Thr Leu Gly Gly Cys Ser Val 115 120 125
75
       Ala Gln Gln Asp Thr Gln Trp Thr Leu Gly Gly Lys Leu Phe Thr Ser
```

```
130
                                  135
                                                         140
       Ala Trp Ile Gln Arg Ser Ala Glu Tyr Gln Ala Leu Cys Ile Gln Ala
145 150 155 160
       Tyr Asn Ile Ala Thr Glu Arg Val Asp Ala Leu Pro Ala Glu Arg Lys
165 170 175
      Gln Gly Asp Arg Pro Tyr Ala Ile Val Thr Asp Ile Asp Glu Thr Ile
180 185 190
       Leu Asp Asn Thr Pro Asn Ser Val Tyr Gln Ala Leu Arg Gly Lys Asp
195 200 205
       Tyr Asp Glu Glu Thr Trp Gly Lys Trp Cys Ala Gln Ala Asp Ala Asp 210 220
10
       Thr Leu Ala Gly Ala Leu Ser Phe Phe Leu His Ala Ala Asn Lys Gly 225 230 235 240
       Ile Glu Val Phe Tyr Val Thr Asn Arg Arg Asp Asn Leu Arg Glu Ala
245 255
15
       Thr Leu Gln Asn Leu Gln Arg Tyr Gly Phe Pro Phe Ala Asp Glu Glu
260 265 270
       His Leu Leu Thr Thr His Gly Pro Ser Asp Lys Glu Pro Arg Arg Leu
275 280 285
       Lys Ile Gln Glu Gln Tyr Glu Ile Val Leu Leu Ile Gly Asp Asn Leu
290 295 300
20
       Gly Asp Phe His His Phe Phe Asn Thr Lys Glu Glu Ser Gly Arg Lys
305 310 315 320
       Gln Ala Leu Gly Leu Thr Ala Gly Glu Phe Gly Arg His Phe Ile Met
325 330 335
25
       Leu Pro Asn Pro Asn Tyr Gly Ser Trp Glu Pro Ala Trp Tyr Gly Gly 340 345 350
                340
       Lys Tyr Pro Pro Leu Pro Glu Arg Asp Lys Ala Leu Lys Gln Leu His
              355
30
       Ser Gln Asn Ser Arg
       (2) INFORMATION FOR SEQ ID NO:288
35
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 426 amino acids(B) TYPE: amino acid
                  (D) TOPOLOGY: linear
40
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
45
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...426
50
           (mi) SEQUENCE DESCRIPTION: SEQ ID NO:288
       Arg Trp Gly Phe Ser Asn Phe Val Ala Met Ser Thr Asn Ile Asp Val
                                               10
55
       Gln Gln Ile Lys Gln Arg Phe Gly Ile Ile Gl; Ser Ser Pro Leu Met 20 \hspace{1cm} 25 \hspace{1cm} 30
       Glu His Ala Ile Arg Val Ala Ala Gln Val Ala Pro Thr Asp Met Ser 35 40 45
       Val Leu Val Thr Gly Glu Ser Gly Ser Gly Lys Glu Phe Phe Pro Gln
50 55 60
60
       Ile Ile His Tyr Tyr Ser Ala Arg Lys His His Ser Tyr Ile Ala Val
65 70 75 80
       Asn Cys Gly Ala Ile Pro Glu Gly Thr Ile Asp Ser Glu Leu Phe Gly
85 90 95
       His Arg Lys Gly Ser Phe Thr Gly Ala Val Ser Asp Arg Lys Gly Tyr
       Phe Glu Glu Ala Ser Gly Gly Thr Ile Phe Leu Asp Glu Val Gly Glu
115 120 125
       Leu Pro Leu Pro Thr Gln Ala Arg Leu Leu Arg Val Leu Glu Thr Gly 130 135
70
                               135
       Glu Phe Ile Pro Val Gly Ala Ser Gln Ser Gln Lys Thr Asp Val Arg
145 150 155 160
       Ile Val Ala Ala Thr Asn Val Asn Leu Lys Glu Ala Val Ala Asn Gly
165 170 175
75
       Lys Phe Arg Glu Asp Leu Phe Phe Arg Leu Asn Thr Val Pro Ile Glu
```

```
180
        Val Pro Ala Leu Arg Het Arg Pro Asp Asp Val Pro Leu Leu Phe Arg
195 200 205
        Arg Phe Ala Ala Asp Ser Ala Glu Lys Tyr Arg Met Pro Pro Leu Arg
210 215 220
  5
        Leu Ser Asp Glu Ala Arg Thr Ile Leu Het Arg Tyr Arg Trp Fro Gly 225 230 235 240
        Asn Val Arg Glu Leu Arg Asn Ile Thr Asp Arg Leu Ser Ile Leu Glu
245 250 255
        Glu Glu Arg Thr Val Ser Ala Glu Thr Ile Thr Arg Tyr Leu Asp Ala 260 265 270
 10
        Glu Gly Het Gln Asp Leu His Pro Val Val Ile Arg Arg Asn Glu Thr
275 280 285
        Thr Glu Ala Asp Lys Gln Ile Pro His Tyr Glu Arg Glu Ile Ile Tyr
290 295 300
        Gln Val Leu Tyr Asp Met Lys Lys Glu Ile Ala Asp Leu Lys Gly Met
305 310 320
       Met Asn Arg Leu Ala His His Glu Gln Pro Ser Trp Pro Val Gly Ser 325 330 335 Asp Val Trp Gly Asn Asp Asp Lys Arg Thr Ala Asp Pro Lys Trp Gly 340 345 350
20
        Val Ser Thr His Lys Ala Pro Ile Ala Asn Ala Ala Glu Pro Val Glu
355 360 365
        Pro Ile Gln Glu Ala Ser Glu Tyr Thr Glu Asp Pro Val Ser Leu Glu
370 375 380
25
        Glu Val Glu Lys Lys Met Ile Ser Leu Ala Leu Glu Arg His Gly Gly 385 390 395 400
        Arg Arg Lys Gln Thr Ala Glu Glu Leu Lys Ile Ser Glu Arg Thr Leu
405 410 415
30
        Tyr Arg Lys Ile Lys Glu Tyr Gly Leu Glu
                    420
        (2) INFORMATION FOR SEQ ID NO:289
35
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 653 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
40
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
45
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...653
50
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:289
       Lys Asn Leu Pro Arg Tyr Ala Pro Ala Phe Gly Arg Asn Arg Lys Arg
                                               10
55
       Glu Gln Pro Asp Pro Asp Asn Asp Asp Ser Arg Gly Arg Ser His Arg 20 25 30
       Gln Ser Gly Arg Ala Val Thr Val Val Cys Arg Arg Ile Gln Thr Pro 35
       Leu Pro His Met Ile Leu Leu Phe Gly Gly Thr Thr Glu Gly Arg Ala 50 60
60
       Ala Ala Arg Val Leu Asp Glu Ala Gly Ser Pro Phe Phe Tyr Ser Thr
65 70 75 80
       Lys Gly Asn Leu Gln Glu Ile Gln Ser Ser His Gly His Arg Leu Thr 85 \hspace{1cm} 90 \hspace{1cm} 95
65
       Gly Ala Het Thr Val Ala Asp Met Val Ser Phe Cys Arg Lys Glu Glu
100 105 110
       Ile Arg Leu Ile Val Asp Ala Ala His Pro Phe Ala Glu Glu Leu His
       Ala Ser Val Ala Glu Ala Ser Glu Gln Thr Gly Ile Pro Val Val Arg
130 135 140
70
                          135
       Tyr Glu Arg Gln Tyr Pro Pro Arg Glu Glu Gly Ile Val Trp Cys Ala
145 150 155 160
       Asn Tyr Asp Thr Ala Ala Glu Arg Met Leu Gly Asp Gly Val Gln Arg
                        165
                                                170
75
       Leu Leu Het Leu Thr Gly Val Asn Thr Ile Pro Lys Leu Ala Ala Phe
```

```
180
                                               185
                                                                        190
       Trp Lys Glu Arg Thr Thr Phe Cys Arg Ile Leu Lys Arg Asp Glu Ser
195 200 205
       Val Ala Leu Ala Glu Lys Asn Gly Phe Pro Ala Glu Arg Ile Val Phe
210 215 220
       Phe Glu Pro His Ala Asp Glu Glu Leu Met Gln Ala Val Arg Pro Asp 225 230 235 240
       Ala Ile Ile Thr Lys Glu Ser Gly Glu Ser Gly Tyr Phe Arg Glu Lys
245 250 255
       Ile Glu Ala Arg Arg Met Gly Ile Arg Ile Tyr Ala Val Arg
260 265 270
10
       Pro Pro Leu Pro Pro Ser Phe Ile Pro Val Gly Gly Pro Val Gly Leu 275 280 285
       Arg Arg Ala Val Glu Arg Leu Val Pro Gly Phe Phe Ser Leu Arg Ser 290 295 300
15
       Gly Phe Thr Thr Gly Thr Thr Ala Thr Ala Ala Val Val Ala Ala Het
305 310 315
       Tyr Arg Leu Het Gly Leu Gly Ser Leu Ala Glu Ala Pro Val Glu Leu 325 330 335
       Pro Ser Gly Glu Tle Val Ser Leu Pro Ile Ala Glu Ile Arg Glu Glu 340 345 350
20
       Glu Asp Ala Val Val Ser Ala Val Leu Lys Asp Ala Gly Asp Asp Pro 355 360 365

Asp Val Thr Asn Gly Met Ala Val Cys Ala Thr Ile Arg Leu Asn Pro 370 380
25
       Glu His Glu Glu Val Arg Phe Leu Gln Gly Glu Gly Val Gly Val Val 385 390 395 400
       Thr Leu Pro Gly Leu Gly Leu Glu Val Gly Gly Pro Ala Ile Asn Leu
405 410 415
30
       Val Pro Arg Arg Met Met Thr Ala Glu Val Arg Arg Leu Tyr Ala Gln
420 425 430
       Gly Gly Val Asp Ile Thr Ile Ser Val Pro Glu Gly Arg Glu Ala Ala
435 440 445
       Thr Gln Thr Phe Asn Pro Arg Leu Gly Ile Arg Asp Gly Ile Ser Ile
450 455 460
35
       The Gly Thr Ser Gly Val Val Lys Pro Phe Ser Ala Glu Ala Phe Val 465 470 480
       Gly Ala Ile Arg Lys Gln Val Gly Ile Ala Thr Ala Leu Gly Ala Asn
485 490 495
40
       His Ile Val Leu Asn Ser Gly Ala Lys Ser Glu Arg Tyr Val Lys Gly 500 505 510
       Ala Tyr Pro Ala Leu Ile Pro Gln Ala Phe Val Gln Tyr Gly Asn Phe
515 520 525
        Val Gly Glu Ser Leu Ser Cys Val Ala Ser Phe Pro Ser Val Arg Ser
530 535 540
45
        Val Thr Val Gly Ile Het Leu Gly Lys Ala Val Lys Leu Ala Glu Gly
545 550 555 560
       Tyr Leu Asp Thr His Ser Lys Lys Val Val Het Asn Arg Asp Phe Leu 565 570 575

His Glu Leu Ala Arg Gln Ala Gly Cys Ser Glu Asp Ile His Ala Ile 580 585 590
50
        Ile Asp Ser Leu Asn Leu Ala Arg Glu Leu Trp Thr Het Pro Ser Ala
595 600 605
        Glu Asp Ser Asp Arg Leu Leu Arg Lys Ile Ala Glu Arg Ser Trp Glu
610 615 620
55
        Thr Cys Arg Pro Ser Val Pro Ser Ala Glu Leu Glu Leu Leu Leu Ile
        Asp Glu Ser Gly Ala Ile Arg Phe Arg Ile Gly Gly Glu
60
        (2) INFORMATION FOR SEQ ID NO:290
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LEMGTH: 451 amino acids
65
                     (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) NOLECULE TYPE: protein
70
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISH: Porphyromonas gingivalis
75
             (ix) FEATURE:
```

```
(A) NAME/KEY: misc_feature
                   (B) LOCATION 1...4\overline{5}
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:290
 5
       Gln Gln Arg Pro His Arg Ile Ile Met Leu Arg Thr Phe Arg Ile Gly
                                                  10
       Gly Ile His Pro Pro Glu Asn Lys Leu Ser Ala Gly Lys Pro Val Glu
                    20
                                           25
                                                                    30
10
       Val Leu Pro Ile Pro Ser Gln Val Val Ile Pro Leu Gly Gln His Ile
35 40 45
       Gly Ala Pro Ala Thr Ala Thr Val Lys Lys Gly Asp Glu Val Lys Val 50 60
       Gly Thr Ile Ile Ala Gln Ala Gly Gly Phe Val Ser Ala Asn Ile His 65 70 75 80
15
       Ser Ser Val Ser Gly Lys Val Leu Lys Ile Asp Asn Val Tyr Asp Ser
85 90 95
       Ser Gly Tyr Pro Lys Pro Ala Val Phe Ile Ser Val Glu Gly Asp Glu
100 105 110
20
       Trp Glu Glu Gly Ile Asp Arg Ser Pro Ala Ile Val Lys Glu Cys Asn 115 120 125
       Leu Asp Ala Lys Glu Ile Val Ala Lys Ile Ser Ala Ala Gly Ile Val
130 135 140
       Gly Leu Gly Gly Ala Thr Phe Pro Thr His Val Lys Leu Ser Pro Pro 145 150 155 160
25
       Pro Gly Asn Lys Ala Glu Ile Leu Ile Ile Asn Ala Val Glu Cys Glu
165 170 175
       Pro Tyr Leu Thr Ser Asp His Val Leu Met Leu Glu His Gly Glu Glu
180 185 190
30
       Ile Met Ile Gly Val Ser Ile Leu Met Lys Ala Ile Gln Val Asn Lys
195 200 205
       Ala Val Ile Gly Val Glu Asn Asn Lys Lys Asp Ala Ile Ala His Leu
210 215 220
       Thr Lys Leu Ala Thr Ala Tyr Pro Gly Ile Glu Val Met Pro Leu Lys 225 230 235 240
35
       Val Gln Tyr Pro Gln Gly Gly Glu Lys Gln Leu Ile Asp Ala Val Ile
245 250 255
       Arg Lys Gln Val Lys Ser Gly Ala Leu Pro Ile Ser Thr Gly Ala Val
260 265 270
       Val Gln Asn Val Gly Thr Val Phe Ala Val Tyr Glu Ala Val Gln Lys
275 280 285
40
       Asn Lys Pro Leu Val Glu Arg Ile Val Thr Val Thr Gly Lys Lys Leu
290 295 300
       Ser Arg Pro Ser Asn Leu Leu Val Arg Ile Gly Thr Pro Ile Ala Ala
305 310 315
45
       Leu Ile Glu Ala Ala Gly Gly Leu Pro Glu Asn Thr Gly Lys Ile Ile
325 330 335
       Gly Gly Gly Pro Met Met Gly Arg Ala Leu Leu Ser Pro Asp Val Pro 340 345 350

Val Thr Lys Gly Ser Ser Gly Val Leu Ile Leu Asp Arg Glu Glu Ala 355 360 365
50
       Val Arg Lys Pro Het Arg Asp Cys Ile Arg Cys Ala Lys Cys Val Gly 370 380
       Val Cys Pro Net Gly Leu Asn Pro Ala Phe Leu Met Arg Asp Thr Leu 385 390 395 400
55
       Tyr Lys Ser Trp Glu Thr Ala Glu Lys Gly Asn Val Val Asp Cys Ile $405$ $410$ $415
       Glu Cys Gly Ser Cys Ser Phe Thr Cys Pro Ala Asn Arg Pro Leu Leu 420 425 430
60
       Asp Tyr Ile Arg Gln Ala Lys Lys Thr Val Met Gly Ile Gln Arg Ala
445 440 445
        Arg Lys Gln
65
        (2) INFORMATION FOR SEQ ID NO:291
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 489 amino acids
                    (B) TYPE: amino acid
70
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
75
```

```
(vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
 5
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...489
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:291
10
      Met Asn Ser Gln Lys Lys Glu Ala Phe Asn Met Lys Arg Ile Gln Leu
                                                 10
       Thr Leu Ile Ala Leu Phe Ala Ala Val Ala Gly Leu Val Ala Gln Asn
20 25 30
       Ala Tyr Glu Gly Val Ile Ser Tyr Lys Ile Ser Leu Asp Lys Thr Gly
35 40
15
       Asn Lys Val Val Leu Asn Gly Ala Ala Asp Het Ser Asn Leu Lys Leu
50 55 60
       Lys Ser Thr Gln Met Ile Ile Val Thr Pro Ile Leu Arg Ser Glu Asp
65 70 80
       Gly Thr Ser Arg Val Glu Phe Pro Ser Val Val Ile Thr Gly Arg Asn
85 90 95
       Arg Thr Lys Ala Leu Lys Arg Glu Ile Ala Phe Ser Ser Ala Leu Pro
100 105 110

Gln Ala Lys His Ala Ala Gln Tyr Ile Arg Arg His Asn Gly Lys Ser
115 120 125
25
       Glu Gln Phe Ala Phe Thr Gly Glu His Ala Tyr Ala Ser Trp Met Het
130 135 140
       Asp Ala Lys Phe Val Val Arg Glu Glu Val Arg Gly Cys Ala Lys Cys
145 150 155 160
30
       Pro Val Gly Leu Ser Ser Asn Ile Val Pro Phe Asp Pro Leu Phe Asn
165 170 175
                       165
       Pro Ala Glu Ala Pro Tyr Leu Leu Ala His Ile Thr Pro Ala Glu Glu
180 185 190
       Val Glu Lys Gln Arg Glu Ser Ser Phe Asp Ala Tyr Ile Asn Phe Lys
195 200 205
35
       Val Asn Lys Ala Asp Val Leu Pro Glu Tyr Arg Asn Asn Lys Ala Glu
210 215 220
       Leu Glu Lys Ile Lys Glu Phe Val Ser Thr Val Lys Ala Asn Pro Asn
225 230 235 240
       Tyr Ser Val Asn Lys Met Ile Ile Glu Gly Phe Ala Ser Pro Glu Ala
245 255
40
       Ser Ile Ala His Asn Lys Ala Leu Ser Glu Arg Arg Ala Lys Arg Leu
260 265 270
       Ala Glu Glu Leu Val Arg Lys Tyr Gly Lys Thr Leu Pro Asn Ile Thr
275 280 285
45
       Thr Glu Phe Gly Gly Glu Asp Trp Lys Gly Leu Lys Leu Ala Ile Glu
290 295 300
       Lys Ser Asp Ile Ala Asp Arg Asp Arg Val Leu Glu Ile Ile Asn Ser
305 310 315 320
       Asp Lys Tyr Ala Asp Asp Asp Ala Arg Glu Gln Ala Leu Lys Gln Leu
325 330 335
50
       Ser Ser Tyr Arg Tyr Ile Leu Asp Gln Ile Tyr Pro Asn Leu Arg Arg
340 345 350
       Asn Thr Ile Thr Met Gly Tyr Ile Val Arg Asp Tyr Thr Leu Glu Glu
355 360 365
55
       Ala Arg Glu Ile Ile Lys Thr Ala Pro Lys Glu Leu Ser Glu Ala Glu
370 375 380
       Het Tyr Arg Val Ala Met Ser Tyr Pro Glu Gly His Gln Glu Arg Leu
385 390 395 400
60
       Phe Ala Leu Asn Thr Thr Leu Lys Tyr Phe Pro Glu Ser Val Thr Gly
405 410 415
       Arg Ile Asn Leu Ala Val Ala Ala Phe Asn Gly Gly Asp Val Gln Gln 420 430
       Ala Ile Ala Leu Leu Ser Pro Ile Gln Thr Glu Lys Gly Val Ser Asn
435 440 445
65
        Ile Leu Gly Ala Ala Tyr Ala Arg Thr Gly Asp Phe Ala Arg Ala Glu
450 455 460
                                 455
        Thr Phe Phe Arg Lys Ala Val Ala Glu Gly Asp Ala Asn Ala Gln Arg
70
        Asn Leu Asp Het Leu Leu Gly Lys Lys
        (2) INFORMATION FOR SEQ ID NO:292
```

(i) SEQUENCE CHARACTERISTICS:

75

```
(A) LENGTH: 384 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 5
            (ii) HOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
10
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...384
15
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:292
       Asp Met Ala Glu Lys Arg Asp Tyr Tyr Glu Val Leu Gly Val Ser Lys
1 10 15
20
       Asn Ala Thr Asp Asp Glu Leu Lys Lys Ala Tyr Arg Lys Lys Ala Ile
20 25 30
       Gln Tyr His Pro Asp Lys Asn Pro Gly Asp Lys Glu Ala Glu Glu His
35 40 45
       Phe Lys Glu Val Ala Glu Ala Tyr Asp Val Leu Ser Asp Pro Gln Lys
50 55 60
25
       Arg Ser Gln Tyr Asp Gln Phe Gly His Ala Gly Leu Gly Gly Ala Ala 65 70 80
       Gly Gly Gly Phe Ser Gly Gly Gly Het Ser Het Glu Asp Ile Phe Ser
85 90 95
30
       Arg Phe Gly Asp Leu Phe Gly Gly Phe Gly Gly Phe Ser 100 105 110
       Asp Met Gly Gly Gly Ser Arg Arg Arg Val Arg Arg Gly Ser Asp Leu
115 120 125
       Arg Val Arg Val Lys Leu Ser Leu Ala Asp Ile Ser Lys Gly Val Glu
130 135 140
35
       Lys Lys Val Lys Val Lys Gln Val Val Cys Ser Lys Cys Arg Gly 145 150 150
       Asp Gly Thr Glu Glu Ala Asn Gly Lys Thr Thr Cys Gln Thr Cys His
165 170 175
       Gly Thr Gly Val Val Thr Arg Val Ser Asn Thr Phe Leu Gly Ala Met
180 185 190
40
       Gln Thr Gln Ser Thr Cys Pro Thr Cys His Gly Glu Gly Glu Ile Ile
195 200 205
       Thr Lys Pro Cys Ser Lys Cys Lys Gly Glu Gly Val Glu Ile Gly Glu
210 215 220
45
       Glu Val Ile Ser Phe His Ile Pro Ala Gly Val Ala Glu Gly Met Gln
225 230 235 240
       Met Ser Val Asn Gly Lys Gly Asn Ala Ala Pro Arg Gly Gly Val Asn 245 250 255
50
       Gly Asp Leu Ile Val Val Ile Ala Glu Glu Pro Asp Pro Asn Leu Ile
260 265 270
       Arg Asn Gly Asn Asp Leu Ile Tyr Asn Leu Leu Ile Ser Val Pro Leu
275 280 285
       Ala Ile Lys Gly Gly Ser Val Glu Val Pro Thr Ile Asp Gly Arg Ala
290 295 300
55
       Lys Ile Arg Ile Glu Ala Gly Thr Gln Pro Gly Lys Met Leu Arg Leu 305 310 315 320

Arg Asn Lys Gly Leu Pro Ser Val Asn Gly Tyr Gly Met Gly Asp Gln 325 335
60
       Leu Val Asn Val Asn Val Tyr Ile Pro Glu Ser Ile Asp Ala Lys Asp 340 345 350
       Glu Gln Ala Ile Ala Ala Het Glu Asn Ser Asp Ser Phe Lys Pro Thr
355 360 365
       Asp Ala Ala Arg Lys Asp Ile Asp Lys Lys Tyr Arg Glu Met Leu Asp 370 380
65
       (2) INFORMATION FOR SEQ ID NO:293
             (i) SEQUENCE CHARACTERISTICS:
70
                   (A) LENGTH: 309 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
75
```

WO 99/29870 PCT/AU98/01023

```
(iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
 5
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...309
10
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:293
       Arg Phe Asp Ser Thr Thr Asn Val Ser Gln Ile Asn Leu Arg Thr Glu
       Met Lys Lys Leu Ile Leu Ala Thr Leu Gly Leu Met Ala Ile Ala Met
15
                                          25
                   20
       Leu Ser Cys Ser Ser Asn Asn Lys Asp Leu Glu Asn Lys Gly Glu Ala 35 \hspace{1cm} 40 \hspace{1cm} 45
       Thr Leu Val Thr Phe Gly Ser Ser Tyr Lys Ala Pro Arg Glu Thr
50 55 60
       Tyr Ala Lys Ile Glu Lys Thr Phe Ala Ala Ala Tyr Pro Asp Gln Arg 65 \hspace{1.5cm} 70 \hspace{1.5cm} 75 \hspace{1.5cm} 80
20
       The Ser Trp Thr Tyr Thr Ser Ser Ile Ile Arg Lys Lys Leu Ala Gln 85 90 95
       Gln Gl; Ile Tyr Ile Asp Ala Pro Asp Glu Ala Leu Glu Lys Leu Ala
100 105 110
25
       Arg Leu Gly Tyr Lys Lys Ile Asn Val Gln Ser Leu His Val Ile Pro 115 120 125
       Gly Arg Glu Tyr Asp Glu Met Ile Asp Phe Val Asn Lys Phe Lys Ala
130 135 140
       Ala His Ser Asp Ile Thr Val Lys Val Gly Ala Pro Leu Phe Asp Thr
145 150 160
30
       Asp Glu Asp Met Arg Glu Val Ala Glu Ile Leu His Lys Arg Phe Gln
165 170 175
                       165
       Gln Thr Ile Glu Lys Gly Glu Ala Ile Val Phe Met Gly His Gly Thr
180 185 190
35
       Glu His Ala Ala Asn Asp Arg Tyr Ala Arg Ile Asn Lys Ile Met Lys
195 200 205
       Asn Tyr Ser Lys Phe Met Ile Val Gly Thr Val Glu Ser Asp Pro Ser
210 215 220
       Ile Asn Asp Val Ile Ala Glu Leu Lys Glu Thr Gly Ala Thr Ala Val
225 230 235
40
       Thr Net Net Pro Leu Met Ser Val Ala Gly Asp His Ala Thr Asn Asp 245 250 255
       Met Ala Gly Asp Glu Asp Asp Ser Trp Lys Thr Leu Leu Thr Asn Ala
260 265 270
45
       Gly Tyr Thr Val Ser Ile Asp Lys Leu Asp Asn Gly Asn Phe Ser Ala
275 280 285
       Leu Gly Asp Ile Glu Glu Ile Arg Asn Ile Trp Leu Lys His Met Lys
290 295 300
           290
50
       Ala Thr Ser Ala Arg
       305
       (2) INFORMATION FOR SEQ ID NO:294
55
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 491 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
60
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
65
                   (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...491
70
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:294
        Arg Arg Asp Ala Arg Arg Gln Leu Val Arg Arg Arg Lys Glu Leu
                                               10
 75
        Ser Gly Thr Ala Asn Thr Arg Cys Arg Arg Arg Val Phe Pro Glu Arg
```

```
Arg Ser Ala Ile Arg Trp Cys Phe Pro Arg Leu Arg Asp Gly Leu Gln 35 40 45
            Pro Ile Arg Phe Gly Ser Leu Leu Ala Asn Pro Gln Gln Leu Ser Arg
50 55 60
            Gly Val Arg Leu Arg Ser Phe Glu Ser Arg Arg Gln Pro Ile Cys Arg
65 70 75 80
            Phe Gln Gly Ile Gln Gly Lys Arg Arg Asp Ile Leu Gln Pro Glu Gly 85 90 95
10
            Asn Ser Leu Leu Arg Thr His Ala His Ala Ser Arg Leu Gly Ala Gly 100 105 110
            Arg Ser Tyr Arg Tyr Leu Ala Pro Arg Thr Thr Ser Arg Pro Pro Thr 115 120 125
            Ala Phe Leu Leu Phe Ala Pro Met Thr Ser Val Ser His Leu Arg Thr
            135 140 140 140 140 140 150 150 150 160 Leu Phe Gly Val Asn Leu Phe Lou Cly Company Co
15
            Leu Phe Gly Val Asn Leu Phe Leu Gly Ser Val Ala Ile Pro Met Ser
165 170 175
20
            Glu Ile Phe Arg His Leu Phe Ser Asp Arg Pro Glu Gly Gly Glu Ala
180 195 190
            Leu Val His Tyr Asn Ile Leu Trp Lys Ser Arg Leu Pro Glu Ala Leu 195 200 205
            Thr Ala Ala Phe Ala Gly Ala Gly Leu Ser Val Ser Gly Leu Gln Net 210 220
25
            Gin Thr Val Phe Arg Asn Pro Leu Ala Gly Pro Ser Val Leu Gly Ile
225 230 235 240
            Ser Ser Gly Ala Ser Leu Gly Val Ala Leu Val Val Leu Leu Ser Gly 245 250 255
            Ser Leu Gly Gly Val Ala Leu Ser Ser Leu Gly Tyr Met Gly Glu Val
260 265 270
30
            Ala Net Asn The Ala Ala Ala Val Gly Ser Leu Ala Val Met Gly Leu
275 280 285
             Ile Val Phe Val Ser Thr Lys Val Arg Ser His Val Thr Leu Leu Ile
290 295 300
35
            Ile Gly Val Met Ile Gly Tyr Val Ala Thr Ala Val Ile Gly Val Phe 305 $310 $315
            Lys Phe Phe Ser Ile Glu Glu Asp Ile Arg Ala Tyr Val Ile Trp Gly 325 330 335
40
             Leu Gly Ser Phe Ser Arg Ala Thr Asp Ser Gln Leu Ser Phe Phe Ala 340 345 350
             Ile Leu Met Leu Ile Phe Ile Pro Ala Gly Met Leu Leu Val Lys Gln 355 360 365
             Leu Asn Leu Leu Leu Gly Glu Ser Tyr Ala Arg Asn Leu Gly Leu
370 375 380
45
             Asn Thr Arg Arg Ala Arg Leu Leu Val Ile Ser Ser Ala Gly Leu Leu
385 390 395 400
             Ile Ala Thr Val Thr Ala Tyr Cys Gly Pro Ile Gly Phe Leu Gly Met 405 415
            Ala Val Pro His Leu Ala Arg Val Ile Phe His Thr Ser Asp His Arg 420 425 430
50
             Ile Leu Met Pro Ala Thr Cys Leu Ile Gly Ser Ala Leu Ala Leu Phe
435
440
445
             Cys Asn Ile Ile Ala Arg Met Pro Gly Phe Glu Gly Ala Leu Pro Val
450 455 460
55
             Asn Ser Val Thr Ala Leu Val Gly Ala Pro Ile Ile Val Thr Val Leu
465 470 475 480
             Phe Arg Arg Arg Phe Lys Glu Glu Thr Asp
                                              485
60
             (2) INFORMATION FOR SEQ ID NO:295
                       (i) SEQUENCE CHARACTERISTICS:
                                 (A) LENGTH: 763 amino acids
65
                                  (B) TYPE: amino acid
                                 (D) TOPOLOGY: linear
                     (ii) MOLECULE TYPE: protein
70
                   (iii) HYPOTHETICAL: YES
                      (vi) ORIGINAL SOURCE:
                                  (A) ORGANISM: Porphyromonas gingivalis
75
                     (ix) FEATURE:
```

(A) NAME/KEY: misc_feature (B) LOCATION 1...763

| _ | | (xi) | SEÇ | DUENC | CE DE | SCRI | PTIC | M: 5 | SEQ I | D HC | 295 | , | | | | |
|------------|------------|----------|------------|------------|------------|------------|------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|
| 5 | | Phe | Leu | Val | Thr | Ile | Met | Arg | Thr | Lչ։s 10 | Thr | lle | Phe | Phe | Ala 15 | Ile |
| | lle | Ser | Phe | Ile 20 | Alā | Leu | Leu | Ser | Ser 25 | | Leu | Ser | Alā | Gln 30 | | Lys |
| 10 | | | 35 | Thr | | | | 40 | Asp | | | | 45 | Glu | | |
| | | 50 | | | | | 55 | | His | | | 60 | | | | |
| 15 | 65 | | | | | 70 | | | Lγε | | 75 | | | | | 80 |
| | | | | | 8.5 | | | | Tyr | 90 | | | | | 95 | |
| 20 | | | | 100 | | | | | Thr 105 | | | | | 110 | | |
| 20 | | | 115 | | | | | 120 | Val | | | | 125 | | | |
| | | 130 | | | | | 135 | | Ala Thr | | | 140 | | | | |
| 25 | 1.45 | | | | | 150 | | | Pro | | 155 | | | | | 160 |
| | | | | | 165 | | | | Tyr | 170 | | | | | 175 | |
| 30 | | | | 180 | | | | | 185 Gly | | | | | 190 | | |
| 30 | | | 195 | | | | | 200 | Glu | | | | 205 | | | |
| | | 210 | | | | | 215 | | Ala | | | 220 | | | | |
| 35 | 225 | | | | | 230 | | | Ala | | 235 | | | | | 240 |
| | _ | | | | 245 | | | | Thr | 250 | | | | | 255 | |
| 40 | | | | 260 | | | | | 265 Phe | | | | | 270 | | |
| | | | 275 | | | | | 280 | Lys | | | | 285 | | | |
| | | 290 | | | | | 295 | | Thr | | | 300 | | | | Gln |
| 4 5 | 305 | | | | | 310 | | | Tyr | Ala | 315 Thr | | | | Ser | 320 |
| | Ser | GI7. | Asn | | | туг | Asn | Lys | Arg | 330 Gln | | Phe | Thr | | 335 Thr | Phe |
| 50 | Ser | Glu | | | | туг | Asp | | 345 Asp | туг | Arg | Ala | | 350 Thr | Ala | Ser |
| | Leu | | | | Туг | Leu | Phe | | Asn | Gly | Leu | His 380 | 365 Thr | Leu | Ser | Phe |
| 55 | Asp 385 | | | ТУr | Asp | Arg 390 | Phe | | Phe | G1?. | Tyr 395 | Leu | T7.r | His | Asp | Lys 400 |
| 55 | Asp | Ser | Ser | Glu | Ser 405 | Leu | | Asn | Asn | Gln 410 | Gly | | Thr | Glu | Gln 415 | Pro |
| | Thr | Phe | Phe | Pro 420 | Gly | Gln | Leu | Arg | Asn 425 | Lγε | | Asp | Gln | Ile 430 | Arg | Туr |
| 60 | The | Ala | Glu 435 | Ala | | Gl y | Val | Phe 440 | Thr | | Pro | Туг | Ala 445 | Gln | Lys | Leu |
| | | 450 |) | | | | 455 | | | | | 460 | | | | Asn |
| 65 | 465 | , | | | | 470 | | | | | 475 | | | | | Gln 480 |
| | | | | | 485 | | | | | 490 | 1 | | | | 495 | |
| | | | | 500 |) | | | | Thr 505 | | | | | 510 | | |
| 70 | | | 515 | i | | | | 520 | • | | | | 525 | | | Asn |
| | | 530 |) | | | | 535 | ; | | | | 540 | | | | Leu |
| 75 | Th: 545 | | Het | G13 | / Ser | His 550 | | Leu | Туг | Leu | 61 y 555 | | Ala | Asp | Leu | Lys 560 |

```
        Pro
        Gln Het
        Ser
        Asp Tyr
        Tyr Ala
        Leu Gly Leu Glu Tyr Asn Gln Gly 575

        Pro
        Ile Ser Phe Ser Ala Thr Val Tyr Asp Asn Glu Leu 580
        Tyr Asp Asn Glu Leu 590
        Sen Leu 590

 5
        Ile Ser Phe Met Asp Ile Pro Thr Ser Pro Glu His Glu Ala Gln Gly 595 600 605
        11e Lys Lys Thr Lys Gln Tyr Ala Asn Ile Gly Lys Ala Arg Ser Arg
610 620
        Gly Leu Asp Val Leu Cys Asp Ala Ser Ile Gly Trp Gly Ile Lys Leu
625 630 635 640
10
        Gly Ala Gly Tyr Ser Leu Val Glu Ala Lys Asn Leu Gln Thr Asp Glu
645 650 655
        Trp Leu Glu Gly Ala Ala Arg His Arg Ala Asn Val His Ala Asp Trp
660 665 670
        Val His Tyr Trp Gly Gln Tyr Arg Leu Gly Val Ser Leu Phe Gly Arg
675 680 685
15
        Ile Gln Ser Glu Arg Tyr Tyr Lys Asp Gly Asn Ala Pro Asp Tyr Thr 690 695 700
        Leu Trp Arg Leu Ala Thr Ser His Arg Phe Ala His Phe Arg His Ile
705 710 715 720
20
       Ile Leu Asp Gly Thr Leu Gly Ile Asp Asn Leu Phe Asp Tyr Val Asp 725 735

Asp Arg Pro Net Gly Val Asn Tyr Ala Thr Val Thr Pro Gly Arg Thr 740 745 750
25
        Phe Phe Ala Gln Ile Ala Ile Arg Phe Asn Asn 755
        (2) INFORMATION FOR SEQ ID NO:296
30
               (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 365 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
35
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
40
                     (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...365
45
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:296
        Cys Cys Ala Arg Leu Arg Asp Thr Ser Pro Arg Lys Lys Ser Phe Arg
1 10 15
                                                      10
50
        Asn Asp Thr Val Phe Pro Tyr Phe Val Ile His Leu Ile Lys His Ile
20 25 30
        Ile Ile Met Thr Asp Asn Lys Gln Arg Asn Ile Val Phe Pro Ala Phe 35 40 45
        Leu Leu Leu Gly Val Ile Ala Val Val Thr Ile Val Gly Phe Phe 50 55
55
        Het Leu Arg Pro Ala Glu Glu Ile Ile Gln Gly Gln Ile Glu Val Thr
65 70 75 80
       Glu Tyr Arg Val Ser Ser Lys Val Pro Gly Arg Ile Lys Glu Leu Arg
85 90
60
        Val Ser Glu Gly Gln Gln Val Gln Ala Gly Asp Thr Leu Ala Val Ile
100 105 110
        Glu Ala Pro Asp Val Ala Ala Lys Met Glu Gln Ala Lys Ala Ala Glu
115 120
        Ala Ala Ala Gln Ala Gln Asn Ala Lys Ala Leu Lys Gly Ala Arg Ser
130 135 140
65
        Glu Gln Ile Gln Ala Ala Tyr Glu Met Trp Gln Lys Ala Gln Ala Gly
145 150 155 160
        Val Ala Ile Ala Thr Lys Thr His Gln Arg Val Gln Asn Leu Tyr Asp
165 170 175
70
       Gln Gly Val Val Pro Ala Gln Lys Leu Asp Glu Ala Thr Ala Gln Arg
180 185 190
        Asp Ala Ala Ile Ala Thr Gln Lys Ala Ala Glu Ala Gln Tyr Asn Het
195 200 205
        Ala Arg Asn Gly Ala Glu Arg Glu Asp Lys Leu Ala Ala Ser Ala Leu
210 215 220
75
```

```
Val Asp Arg Ala Arg Gly Ala Val Ala Glu Val Glu Ser Tyr Ile Asn 225 230 235 240
       Glu Thr Tyr Leu Ile Ala Pro Arg Ala Gly Glu Val Ser Glu Ile Phe
                                              250
 5
       Pro Lys Ala Gly Glu Leu Val Gly Thr Gly Ala Pro Ile Met Asn Ile
265 270
       Ala Glu Met Gly Asp Met Trp Ala Ser Phe Ala Val Arg Glu Asp Phe
275 280 285
       Leu Ser Ser Met Thr Met Gly Ala Val Leu Glu Thr Val Val Pro Ala
290 295 300
10
       Leu Asn Glu Glu Lys Val Arg Phe Lys Ile Thr Phe Ile Lys Asn Met 305 310 315 320
       Gly Thr Tyr Ala Ala Trp Lys Ala Thr Lys Thr Thr Gly Gln Tyr Asp 325 \hspace{1.5cm} 330 \hspace{1.5cm} 335
       Leu Lys Thr Phe Glu Val Lys Ala Thr Leu Ala Asp Lys Asp Lys Ala 340 345 350
15
       Gln Lys Leu Arg Pro Gly Met Ser Val Ile Ile Arg Lys
355 360 365
20
       (2) INFORMATION FOR SEQ ID NO: 297
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 320 amino acids
                   (B) TYPE: amino acid
25
                   (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
35
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...320
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:297
40
       Pro Gln Ser Ser Pro Asp Arg Arg Ser Phe Gln Asn Val Met Asn Lys
       Tyr His Ser Gln Ser Val Leu Glu Val Gly Lys Ile Gly Ile Val Ile
20 25 30
       Ile Phe Ala Pro Ile Val Arg Asn Val His Gln Gln Pro Pro Phe Leu
35 40 45
45
       Ser His Lys Ser Ile Met Arg Ile Val Ser Asn Phe Leu Phe Val Ser
50 55
       Phe Ser Val Leu Leu Phe Ala Ser Cys Arg Ser Gln Arg Glu Lys Val
65 70 75 80
       Val Tyr Leu Gln Asp Ile Gln Thr Phe Asn Arg Glu Ile Ile Ala Lys
85 90
50
       Pro Tyr Asp Val Lys Ile Glu Lys Asp Asp Val Leu Asn Ile Leu Val
100 105
       Ser Ser Arg Asp Pro Glu Leu Ser Thr Pro Tyr Asn Gln Val Leu Thr
55
       Thr Arg Ala Leu Ala Arg Asn Gly Tyr Gly Thr Asn Ser Asn Glu Gly
130 135 140
       Phe Leu Val Asp Ser Lys Gly Tyr Ile Asn Tyr Pro Ile Leu Gly Gln
145 150 160
60
       Ile Tyr Val Glu Gly Leu Thr Arg Thr Glu Leu Glu Lys Glu Ile Gln 165 170 175
       Lys Arg Ile Ile Ser Ser Gly Phe Ile Lys Asp Pro Thr Val Thr Val 180 185 190
       Gln Leu Gln Asn Phe Lys Val Ser Val Leu Gly Glu Val Asn His Pro
195 200 205
65
       Gly Ser Het Ser Val Lys Gly Glu Arg Ile Thr Leu Leu Glu Ala Ile
210 215 220
       Gly Met Ala Gly Asp Leu Thr Ile Tyr Gly Arg Arg Asp Arg Val Phe 225 230 235 240
70
       Val Ile Arg Glu Thr Asp Gly His Arg Glu Val Phe Gln Thr Asp Leu
245 250 255
       Arg Lys Ala Asp Leu Leu Ala Ser Pro Val Tyr Tyr Leu His Gln Asn 260 265 270
       Asp Val Ile Tyr Val Glu Pro Asn Asp Lys Lys Thr Gln Met Ser Glu
75
                                       280
```

```
Ile Asn Gln Asn Asn Asn Val Asn Val Trp Leu Ser Val Thr Ser Thr 290 295 300
            290 295
                                                    300
        Leu Val Ser Ile Ser Thr Leu Thr Ile Thr Ile Ile Asp Lys Thr Lys
 5
        (2) INFORMATION FOR SEQ ID NO:298
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 582 amino acids
10
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
15
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivali\varepsilon
20
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...582
            (:i) SEQUENCE DESCRIPTION: SEQ ID NO:298
25
       Ser Lys Ile Val Leu Arg Lys Phe Cys Thr Leu Ala Arg Met Lys Lys 1 5 10 15
       Thr Asn Leu Phe Leu Ser Leu Leu Val Ile Phe Ile Thr Gl_Y Ser Phe 20 25 30
30
       Het Thr Ala Cys Ala Gln Lys Ser Lys Thr Asn Lys Leu Thr Glu Glu 35 40
       Ile Ile Gly Asn Val Arg Leu Tyr Phe Val Asp Thr Ile Ser Ile Lys 65 70 75 80
35
       His Net Thr Arg Arg Gly Ile Asp Ala Net Leu Gly Gly Leu Asp Pro 85 \hspace{1.5cm} 90 \hspace{1.5cm} 95
       Tyr Thr Glu Tyr Ile Pro Tyr Glu Glu Het Asp Glu Leu Lys Leu Met 100 105 110
40
       Thr Thr Gly Glu Tyr Ala Gly Val Gly Ala Ile Ile Ser Gln Arg Pro
115 120 125
       Asp Ser Ala Val Ile Ile Gln Arg Pro Met Glu Gly Met Pro Ala Asp
130 135 140
       Glu Ala Gly Leu Ile Ala Gly Asp Arg Ile Leu Thr Ile Asp Gly Lys
145 150 156
45
       Asp Phe Arg Lys Ser Thr Thr Pro Lys Val Ser Gln Ala Leu Lys Gly
165 176 175
       Ile Ala Gly Thr Val Ala Lys Val Thr Val Het Arg Tyr Gly Glu Thr
50
       Lys Pro Arg Thr Phe Ser Val Lys Arg Gln Lys Val Ile Het Asn Ser
195 200 205
       Val Thr Tyr Ser Gly Met Leu Asp Gly Ser Ile Gly Tyr Ile Arg Leu
210 215 220
       Asn Asn Phe Thr Asp Lys Ser Ala Glu Glu Val Arg Thr Ala Leu Leu
225 230 235 240
55
       Asp Leu Arg Asp Lys Gln Gly Ala Lys Gl; Leu Ile Leu Asp Leu Arg 245 255
       Gly Asn Gly Gly Leu Met Gln Ala Ala Ile Glu Ile Val Asn Leu 260 265 270

Phe Val Pro Lys Gly Lys Glu Val Val Thr Thr Lys Gly Arg Ile Ala 275 280 285

Cly Cor Ala Sar Val Pho Arg Thr Low Thr Gly Pro Ile Asn Thr Lys
60
       Glu Ser Ala Ser Val Phe Arg Thr Leu Thr Glu Pro Ile Asp Thr Lys
290 295 300
       Leu Pro Ile Val Val Leu Ile Asp Gly Gln Ser Ala Ser Ser Glu 305 310 315 320
65
       Ile Val Ala Gly Ala Leu Gln Asp Net Asp Arg Ala Val Leu Met Gly 325 330 335
       Gln Lys Ser Tyr Gly Lys Gly Leu Val Gln Thr Thr Arg Gln Leu Pro 340 345 350
       Tyr Asn Gly Val Ile Lys Leu Thr Thr Ala Lys Tyr Tyr Ile Pro Ser 355 360 365
70
       Gly Arg Cys Ile Gln Arg Leu Asp Tyr Ser Arg Thr Asn Arg Thr Gly
370 375 380
       Met Ala Thr Ala Ile Pro Asp Ser Leu His Lys Ile Phe Tyr Thr Ala
385 390 395 400
75
```

```
Val Lys Gln Asp Thr Ala Ala Thr Leu Leu Tyr Tyr Met Ala Ile Asn
420 425 430
      Asn Asp Val Phe Asp Phe Val Thr Gly Tyr Val Leu Lys His Lys Thr
435 440 445
       Ile Ala Lys Pro Glu Asp Phe Ser Ile Thr Asn Glu Asp Tyr Ala Ala
450 455 460
       Phe Cys Lys Het Het Glu Glu Lys Lys Phe Asp Tyr Asp Arg Gln Ser
465 470 475 480
10
       Gly Lys Met Leu Asp Lys Leu Glu Glu Leu Ala Lys Ile Glu Gly Tyr
485 490 495
       Leu Pro Glu Ala Asn Ser Glu Leu Lys Ala Leu Arg Glu Lys Leu Lys
500 505 510
       Pro Asn Leu Ser Arg Asp Leu Leu Arg Phe Lys Lys Glu Ile Thr Asn 515 525
15
       Tyr Leu Asn Asn Glu Ile Val Thr Arg Tyr Tyr Tyr Glu Arg Gly Ser
530 540
       Ile Arg Gln Ser Leu Pro Glu Asp Lys Val Val Lys Glu Ala Ile Lys 545 550 560
20
       Leu Leu Lys Asp His Pro Glu Gln Ile Arg Gln Ile Leu Ala Ala Pro
                      565
       Lys Ala Glu Asn Lys Gly
                    580
25
       (2) INFORMATION FOR SEQ ID NO:299
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 985 amino acids
30
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
40
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...985
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:299
45
       Val Thr Asp Lys Het Gln Asn Lys Gly Phe Val Ile Val Ile Thr Ser
                                               10
      Ala Leu Ala Ile Ile Cys Ala Phe Tyr Leu Ser Phe Ser Phe Val Thr
20 25 30
      Asn Arg Tyr Glu Lys Lys Ala Lys Ala Met Gly Asp Val Ala Gly Met 35 40
      Ala Tyr Leu Asp Ser Met Ser Asn Glu Lys Val Trp Phe Gly Tyr Thr 50 60
      Leu Lys Glu Ala Gln Ala Gln Gln Ile Gly Leu Gly Leu Asp Leu Lys 65 70 75 80
55
      Gly Gly Met Asn Val Ile Leu Lys Leu Asn Ala Ser Asp Leu Leu Arg
                                            90
      Asn Leu Ser Asn Lys Ser Leu Asp Pro Asn Phe Asn Lys Ala Leu Glu
100 105 110
60
      Asn Ala Ala Lys Ser Thr Glu Gln Ser Asp Phe Ile Asp Ile Phe Val
115 120 125
      Lys Glu Tyr Arg Lys Leu Asp Pro Asn Gly Arg Leu Ala Val Ile Phe
130 140
      Gly Ser Gly Asp Leu Arg Asp Gln Ile Thr Ala Lys Ser Thr Asp Ala
145 155 160
65
      Asp Val Val Arg Leu Leu Lys Glu Lys Tyr Asn Ser Ala Val Glu Ala
165 170 175
      Ser Phe Asn Val Leu Arg Ala Arg Ile Asp Ala Phe Gly Val Val Ala
180 185 190
70
       Pro Asn Leu Gln Arg Leu Glu Gly Gln Gly Arg Ile Leu Val Glu Leu
195 200 205
       Pro Gly Val Lys Asp Pro Glu Arg Val Arg Thr Leu Leu Gln Arg Ser
210 215 220
       Ala Asn Leu Gln Phe Trp Arg Thr Tyr Lys Phe Glu Glu Val Ser Gly
225 230 235 240
75
```

Ala Gly Arg Arg Val Glu Asp Ala Gly Gly Ile Leu Pro Asp Ile Glu 405 415

410

405

| | Asp | Leu | ı Ils | : Ala | Ala 245 | Asr | Asp | Arç | l Fer | . Ser 250 | | Leu | Ala | Met | Asr 255 | Asn |
|------------|------------------------|------------|--------------|------------|------------|------------|------------|------------|------------|--------------|------------|------------|------------|------------|------------|------------|
| | Thr | Asp | Ala | Thr 260 | Pro | Glu | Thr | G1. | Pro 265 | Ala | Thr | Thr | Asp | Ser 270 | Val | Ala |
| 5 | Ala | Thr | Ala 275 | Asp | Ser | Ala | Ala | Val 280 | Glr | Ala | Val | Ala | Asp 285 | Ser | Ala | Thr |
| | Val | Ala 290 | Gln | | Glu | Ala | Lys 295 | Asp | Ala | Thr | Arg | Lys 300 | Asp | Ala | Leu | Phe |
| 10 | 305 | | | | | 310 |) | | | | 315 | Val | Val | | | Ala 320 |
| | | | | Asn | 325 | | | | | 330 | Het | Leu | | | 335 | His |
| | | | | Val 310 | | | | | 345 | | | | | 350 | Ala | Lys |
| 15 | | | 355 | | | | | 360 | | | | | 365 | | | |
| | | 370 | | Arg | | | 375 | | | | | 380 | | | | |
| 20 | 362 | | | Ser | | 390 | | | | | 395 | | | | | 400 |
| | | | | Val | 405 | | | | | 410 | | | | | 415 | |
| 25 | | | | Thr 420 | | | | | 425 | | | | | 430 | | |
| 20 | | | 435 | | | | | 440 | | | | | 445 | | | |
| | | 450 | | Gln Val | | | 455 | | | | | 460 | | | | |
| 30 | 465 | | | Val Asn | | 4/0 | | | | | 475 | | | | | 480 |
| | | | | Leu | 485 | | | | | 490 | | | | | 495 | |
| 35 | | | | 500 Ala | | | | | 505 | | | | | 510 | | |
| | | | $^{\rm oto}$ | Asn | | | | 520 | | | | | 525 | | | |
| | | 220 | | Thr | | | 535 | | | | | 540 | | | | |
| 4 0 | 545 | | | Asp | | 550 | | | | | 555 | | | | | 560 |
| | | | | Gly | 565 | | | | | 570 | | | | | 575 | |
| 4 5 | | | Phe | 580 Ser | | | | | 585 | | | | | 590 | | |
| | | lle | 595 | Leu | | | Tyr | 600 | | | | | 605 | | | |
| 50 | Thr | OTO | | Ile | | Gly | 615 | | | | Phe | 620 | | | | |
| 30 | 625 Leu | Thr | Arg | Ile | Val | 630 Phe | Glu | Lys | Leu | Ala | 635 Lys | r).e | Gly | Arg | Leu | 640 Asp |
| | ${\rm Ly} \varepsilon$ | Ile | Thr | Phe | 645 Thr | Thr | Ser | Ile | Thr | 650 Arg | Asn | Leu | Leu | Val | 655 Asn | Pro |
| 55 | Ser | туг | Asn 675 | 660 Ile | Leu | Gly | Lys | Arg | 665 Lys | Thr | Gly | Phe | | 670 Ile | Pro | Val |
| | Ile | Ile 690 | | Vāl | Leu | Glу | Leu 695 | 680 Ile | Ala | Ser | Phe | | 685 Ile | Gly | Leu | Asn |
| 60 | Arg 705 | | Ile | Glu | Phe | Ser 710 | Gly | Gly | Arg | Asn | T;r 715 | 700 Val | Val | Lys | | |
| | Gln | Pro | Val | Ser | Ser 725 | Glu | Ala | Val | Arg | Ser 730 | Ala | Leu | Ser | Ser | Pro | 720 Leu |
| | Gln | Glu | Lys | Val 740 | | Val | Thr | Ser | Ile 745 | Gly | Thr | Glu | | Thr 750 | 735 Glu | Val |
| 65 | Arg | lle | Ser 755 | Thr | Asn | Туг | r2.e | Il∈ 760 | Gln | Glu | Glu | Ser | Glu 765 | Glu | Thr | Glu |
| | Ala | Glu 770 | Ile | Thr | Asp | Lys | Leu 775 | Tyr | Glń | Ser | Leu | Lys 780 | Gly | Phe | Туг | Thr |
| 7 0 | Gln 4 | | | | | 790 | | | | | 795 | Ile | | | | BUU |
| | Lys ' | | | | Ser 805 | Met | | | | Ile 810 | Thr | | | | 215 | Trp |
| | Ala ' | | | 020 | | | | | 825 | Ala | | | | Leu 830 | Ile | |
| 75 | Phe 2 | Arg | Asp | | Ser | Phe | Ser | Ala | Gly | Val | Phe | Val | Ser | Val | Ala | Ala |

WO 99/29870 PCT/AU98/01023

```
Thr Thr Phe Cys Ile Ile Ala Leu Tyr Ala Leu Leu Trp Lys Ile Leu 850 855 860
       Pro Phe Thr Net Glu Ile Asp Gln Asn Phe Ile Ala Ala Ile Leu Ala
865 870 875 880
 5
       Ile Ile Gly Tyr Ser Leu Asn Asp Thr Val Val Phe Asp Arg Ile
885 890 895
       Arg Glu Thr Met Lys Leu Tyr Pro Asn Arg Asp Arg Tyr Gln Val Ile
900 910
10
       Asn Asp Ala Leu Asn Ser Thr Leu Gly Arg Thr Leu Asn Thr Ser Leu 915 920 925
       Thr Thr Phe Ile Val Het Leu Val Ile Phe Ile Phe Gly Gly Ala Thr 930 935 940
       Net Arg Ser Phe Thr Phe Ser Ile Leu Leu Gly Ile Val Ile Gly Thr
945 950 955 960
15
       Tyr Ser Thr Leu Phe Val Ala Thr Pro Leu Ala Tyr Glu Ile Gln Lys
                       965
       Arg Lys Leu Asn Lys Ala Ala Lys Lys
20
       (2) INFORMATION FOR SEQ ID NO:300
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1046 amino acids(B) TYPE: amino acid
25
                  (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
30
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
35
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1046
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:300
40
       Asp Phe Pro Trp Val Arg Val Lys Pro Glu Lys Lys Arg Lys Gln His
       Asn Ser Asn Asn Phe Lys Phe Asn Ala Lys Glu Lys Ser Met Lys Arg
                                          25
45
       Met Thr Leu Phe Phe Leu Cys Leu Leu Thr Ser Ile Gly Trp Ala Met 35 40 45
       Ala Gln Asn Arg Thr Val Lys Gly Thr Val Ile Ser Ser Glu Asp Asn 50 55
       Glu Pro Leu Ile Gly Ala Asn Val Val Val Val Gly Asn Thr Thr Ile
65 70 80
50
       Gly Ala Ala Thr Asp Leu Asp Gly Asn Phe Thr Leu Ser Val Pro Ala
                       85
                                             90
       Asn Ala Lys Met Leu Arg Val Ser Tyr Ser Gly Met Thr Thr Lys Glu
100 105 110
55
       Val Ala Ile Ala Asn Val Met Lys Ile Val Leu Asp Pro Asp Ser Lys 115 120 125
       Val Leu Glu Gln Val Val Val Leu Gly Tyr Gly Thr Gly Gln Lys Leu
130 135 140
       Ser Thr Val Ser Gly Ser Val Ala Lys Val Ser Ser Glu Lys Leu Ala
145 150 155 160
60
       Glu Lys Pro Val Ala Asn Ile Met Asp Ala Leu Gln Gly Gln Val Ala
165 170 175
       Gly Met Gln Val Net Thr Thr Ser Gly Asp Pro Thr Ala Val Ala Ser
180 190
65
       Val Glu 11e His Gly Thr Gly Ser Leu Gly Ala Ser Ser Ala Pro Leu
195 200 205
       Tyr Ile Val Asp Gly Met Gln Thr Ser Leu Asp Val Val Ala Thr Met 210 215 220
       Asn Pro Asn Asp Phe Glu Ser Met Ser Val Leu Lys Asp Ala Ser Ala
225 230 235 240
70
       Thr Ser Ile Tyr Gly Ala Arg Ala Ala Asn Gly Val Val Phe Ile Gln
245 250 255
       Thr Lys Lys Gly Lys Met Ser Glu Arg Gly Arg Ile Thr Phe Asn Ala
                    260
                                          265
75
       Ser Tyr Gly Ile Ser Gln Ile Leu Asn Thr Lys Pro Leu Asp Asn Het
```

| | | | 275 | | | | | 280 | | | | | 285 | | | |
|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Het | Thr 290 | Gly | Asp | Glu | Leu | Leu 295 | Asp | Phe | Gln | Val | Lys 300 | Ala | Gly | Phe | Trp |
| 5 | G1 y 305 | | Asn | Gln | Thr | Val 310 | | Lys | Val | Lys | Λεp 315 | | Ile | Leu | Ala | G1; 320 |
| Ü | | Glu | Asp | Leu | Tyr 325 | | Asn | Tyr | Asp | Ser 330 | | Lys | Asp | Glu | Туг 335 | |
| | rl.a | Thr | Leu | Phe 340 | | Vāl | Asp | Phe | Asn 345 | His | Asp | Ala | Asp | Trp 350 | | Γλε |
| 10 | Ala | Leu | Phe 355 | r).2 | Thr | Ala | Pro | Thr 360 | Ser | Gln | Gly | Asp | Ile 365 | Ser | Phe | Ser |
| | G1?. | Gly 370 | Ser | Gln | G1? | Thr | Ser 375 | Tyr | Tyr | Ala | Ser | 11e 380 | Glγ | Туг | Phe | Asp |
| 15 | Gln 385 | Glu | Gly | Net | Ala | Arg 390 | Glu | Pro | Ala | Asn | Phe 395 | Γλε | Arg | Tyr | Ser | Gly 400 |
| | Arg | Leu | Asn | Phe | Glu 405 | Ser | Arg | Ile | Asn | Glu 410 | Trp | Leu | Lys | Val | Gly 415 | Ala |
| | Asn | Leu | Ser | Gly 420 | Ala | Ile | Ala | Asn | Arg 425 | Arg | Ser | Ala | Asp | Tyr 430 | Phe | G1 7. |
| 20 | Γλε | Tyr | Tyr 435 | Het | Gly | Ser | Glγ | Thr 440 | Phe | Gly | Vāl | Leu | Thr 445 | Met | Pro | Arg |
| | Туг | Tyr 450 | Asn | Pro | Phe | Asp | Val 455 | Asn | Gly | Asp | Leu | Ala 460 | Asp | Vāl | Туг | Туг |
| 25 | 465 | | | | | 470 | | | | Thr | 475 | | | | | 480 |
| | | | | | 485 | | | | | G1n 490 | | | | | 495 | |
| •• | | | | 500 | | | | _ | 505 | Thr | | | | 510 | | |
| 30 | | | 515 | | | | | 520 | | Ser | | | 525 | | | |
| | | 530 | | | | | 535 | | | Arg | | 540 | | | | |
| 35 | Asp 545 | Val | Ser | Lys | Ser | Phe 550 | Thr | Asn | Thr | Ala | Glu 555 | Tyr | Lys | Phe | Ser | 11e 560 |
| | Asp | Glu | Ly:s | His | Asp 565 | Leu | Thr | Ala | Leu | Met 570 | Glу | His | Glu | Tyr | Ile 575 | Glu |
| 40 | Ту·r | Glu | Gly | Asp 580 | Val | Ile | Gly | Ala | Ser 585 | Ser | Lys | Gly | Phe | G1u 590 | Ser | Asp |
| 4 0 | | | 595 | | | | | 600 | | Thr | | | 605 | | | |
| | | 610 | | | | | 615 | | | Tyr | | 620 | | | | |
| 45 | 625 | Asn | Tyr | GΙγ | Phe | 630 | rλs | Trp | Met | Tyr | 11e 635 | Asp | Phe | Ser | Val | Arg 640 |
| | | | | | 645 | | | _ | | Asn 650 | | | | | 655 | |
| . | Tyr | Ser | Val | 660 Gl7 | Gly | Het | Phe | Asp | Il≘ 665 | Tyr | Asn | r?s | Phe | 11e 670 | Gln | Glu |
| 50 | | | 675 | | | | | 680 | | Lys | | | 685 | - | | |
| | Gly | Asn 690 | Ser | Glu | Ile | Gly | Asn 695 | Tyr | Asn | His | Gln | Ala 700 | Leu | Vāl | Thr | Val |
| 55 | 705 | | | | | 710 | | | _ | Leu | 715 | | | | | 720 |
| | | | | | 725 | | | | | Ser 730 | | | | | 735 | |
| 00 | Ala | Ala | Gly | Ala 740 | Phe | Asn | Asn | Arg | Leu 745 | Ser | Ala | Glu | Val | Asp 750 | Phe | Тут |
| 60 | | | 755 | | | - | | 760 | | Asp | | | 765 | | - | |
| | | 770 | | | | | 775 | | | Val | | 780 | | | | |
| 65 | Gly 785 | Val | Asp | Leu | Ser | Leu 790 | Lys | Gly | Thr | Ile | Tyr 795 | Gln | Asn | Lys | Asp | Trp 800 |
| | | | | | 805 | | | | | Tyr 810 | | | | | 815 | |
| 70 | | | | 820 | | | | | 825 | Het | | | | 830 | | |
| 70 | | | 835 | | | | | 840 | | Phe | - | | 845 | | _ | |
| | | 850 | | | | | 855 | | | Leu | | 860 | | | | |
| 75 | Val 865 | Asp | Ala | Asp | Gly | Asn 870 | ГÀЗ | Val | Thr | Thr | Ser 875 | Gln | Tyr | Ser | Ala | Asp 880 |
| | | | | | | | | | | | | | | | | |

```
Leu Glu Thr Arg Ile Asp Lys Ser Val Thr Pro Pro Ile Thr Gly Gly 885 890 895
       Phe Ser Leu Gly Ala Ser Trp Lys Gly Leu Ser Leu Asp Ala Asp Phe
                     900
                                           905
 5
       Ala Tyr Ile Val Gly Lys Trp Het Ile Asn Asn Asp Arg Tyr Phe Thr
915 920 925
       Glu Asn Ala Gly Gly Leu Met Gln Leu Asn Lys Asp Lys Met Leu Leu
930 935 940
       Asn Ala Trp Thr Glu Asp Asn Lys Glu Thr Asp Val Pro Lys Leu Gly 945 950 950 960
10
       Gln Ser Pro Gln Phe Asp Thr His Leu Leu Glu Asn Ala Ser Phe Leu
965 970 975
       Arg Leu Lys Asn Leu Lys Leu Thr Tyr Val Leu Pro Asn Ser Leu Phe 980 985 990

Ala Gly Gln Asn Val Ile Gly Gly Ala Arg Val Tyr Leu Met Ala Arg 995 1000 1005

Asn Leu Leu Thr Val Thr Lys Tyr Lys Gly Phe Asp Pro Glu Ala Gly 1010 1015 1020
15
       Gly Asn Val Gly Lys Asn Gln Tyr Pro Asn Ser Lys Gln Tyr Val Ala
1025 1030 1035
20
       Gly Ile Gln Leu Ser Phe
       (2) INFORMATION FOR SEQ ID NO:301
25
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 869 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
30
            (ii) HOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
35
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
40
                   (B) LOCATION 1...869
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:301
       Trp His Arg Asn Ile Phe Ile Phe Ala Ser Thr Phe Ser Pro Lys Asn
45
       Met Leu Pro Leu Pro Tyr Arg Tyr Ala Lys Thr Glu His Leu Phe Leu
20 25 30
       Ala Lys Gly Tyr Cys Lys Asn Pro Ile Thr Asn Ile Ile Ile Leu Phe 35 40
50
       Met Lys Lys Lys Asn Phe Leu Leu Gly Ile Phe Val Ala Leu Leu 50 \phantom{\bigg|}55\phantom{\bigg|}
       Thr Phe Ile Gly Ser Net Gln Ala Gln Gln Ala Lys Asp Tyr Phe Asn 65 70 75 80
       Phe Asp Glu Arg Gly Glu Ala Tyr Phe Ser Phe Lys Val Pro Asp Arg
85 90 95
55
       Ala Val Leu Gln Glu Leu Ala Leu Ile Met Ser Ile Asp Glu Phe Asp
                     100
                                           105
       Pro Val Thr Asn Glu Ala Ile Ala Tyr Ala Ser Glu Glu Glu Phe Glu 115 120 125
               115 120
60
       Ala Phe Leu Arg Tyr Gly Leu Lys Pro Thr Phe Leu Thr Pro Pro Ser
130 135 140
       Het Gln Arg Ala Val Glu Met Phe Asp Tyr Arg Ser Gly Glu Lys Tyr
145 150 155 160
       Glu Trp Asn Ala Tyr Pro Thr Tyr Glu Ala Tyr Ile Ser Met Met Glu 165 170 175
65
       Glu Phe Gln Thr Lys Tyr Pro Ser Leu Cys Thr Thr Ser Val Ile Gly
180 185 190
       Lys Ser Val Lys Asp Arg Lys Leu Met Ile Cys Lys Leu Thr Ser Ser 195 200 205
70
       Ala Asn Thr Gly Lys Lys Pro Arg Val Leu Tyr Thr Ser Thr Met His
210 215 220
       Gly Asp Glu Thr Thr Gly Tyr Val Val Leu Leu Arg Leu Ile Asp His
225 230 235 240
       Leu Leu Ser Asn Tyr Glu Ser Asp Pro Arg Ile Lys Asn Ile Leu Asp
245 250 255
75
```

| | | | | 260 |) | | | | 265 |) | | | | 270 |) | туг |
|----------------|------------|-----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| - | | | 275 |) | | | | 280 | Gl) | ' Ala | | | 285 | Asn | Ala | a Asn |
| 5 | | 290 |) | | | | 295 |) | | | | 300 | Ala | Gl y | | His |
| | 305 | , | | | | 310 | | | | | 315 | | | | | Leu 320 |
| 10 | | | | | 325 | | | | | 330 |) | | | | 335 | Thr |
| | | | | 340 |) | | | | 345 | | | | | 350 | | Asp |
| 15 | | | 355 | | | | | 360 | • | | | | 365 | | | Gln |
| | | 370 | • | | | | 375 | | | | | 380 | | | | Ile |
| | 385 | | | | | 390 | | | | | 395 | | | | | 400 Thr |
| 20 | | | | | 405 | | | | | 410 | | | | | 415 | |
| | | | | 420 | | | | | 425 | | | | | 430 | | Gly |
| 25 | | | 435 | | | | | 440 | | | | | 445 | | | Leu |
| | | 450 | | | | | 455 | | | | | 460 | | | | Leu |
| | 465 | | | | | 470 | | | | | 475 Gly | | | | | 480 |
| 30 | | | | | 485 | | | | | 490 | | | | | 495 | |
| | | | | 500 | | | | | 505 | | Leu | | | 510 | | |
| 35 | | | 515 | | | | | 520 | | | Net | | 525 | | | |
| | | 530 | | | | | 535 | | | | Asn | 540 | | | | |
| •• | 545 | | | | | 550 | | | | | 555 Het | | | | | 560 |
| 40 | | | | | 565 | | | | | 570 | Туг | | | | 575 | |
| | | | | 580 | | | | | 585 | | Thr | | | 590 | | |
| 4 5 | | Val | 595 | | | | | 600 | | | Phe | | 605 | | | |
| | Glu | 610 | | | | Gln | 615 | | | | Gln | 620 | | | | |
| 50 | 625 | | | | Val | 630 | | | | | 635 Gly | | | | | 640 |
| 50 | | | | Asn | 645 | | | | Tyr | 650 | Γλ.ε | | | | 655 | |
| | | | Leu | 660 | | | | Ala | 665 | | Glu | | Val | 670 | | |
| 55 | | Lys | Tyr | | | | Lys | 680 | | | Val | Pro | 685 | | | |
| | | 690 | | | | Pro | 695 | | | | Lys | 700 | | | | Thr |
| 60 | | Lys | Asp | Leu | Ser 725 | 710 Thr | Asn | Asn | Pro | | 715 Ser | Trp | Leu | Trp | | 720 Phe |
| ~ - | Glu | Gly | Gly | Ser 740 | | Ala | Thr | Ser | Thr | 730 Glu | Gln | Asn | Pro | | 735 Val | Thr |
| | Tyr | Asn | Glu 755 | | Gly | L7.2 | Туг | Asp 760 | 745 Val | Gln | Leu | Thr | | 750 Thr | Asn | Glu |
| 65 | G17. | G1; | | Asn | Val | Lys | Lys 775 | Ala | Glu | Asp | туr | | 765 Glu | Val | Ile | Leu |
| | Asp 785 | | Ser | Val | Glu | | | Val | Ala | Gln | Thr 795 | 780 Gly | Ile | Val | Ile | |
| 70 | | Gln | Asn | Gl y | | | Gln | Ile | Leu | Ile 810 | Glu | Ala | Asn | | Ala 815 | 800 Ile |
| | Γλε | Ala | Ile | Val 820 | | ТУг | Asp | Ile | Asn 825 | Gly | Arg | Väl | | Leu 830 | F7.2 | Thr |
| | Thr | Pro | Asn 835 | | Leu | Arg | Ser | Thr 840 | Val | Asp | Leu | | 11e 845 | Leu | Pro | Glu |
| 75 | Gly | lle | | Thr | Ile | Asn | Ile | Lys | Thr | Glu | Lys | Ser | Ala | Arg | Thr | Glu |

| | 850 Lys Ile His Ile Gly 865 | 855 | 860 | | | | | | |
|------------|--|-----------------------------------|------------------------------|--|--|--|--|--|--|
| 5 | (2) INFORMATION FOR SEQ | ID NO:302 | | | | | | | |
| 10 | (i) SEQUENCE CHARA (A) LENGTH: 1 (B) TYPE: ami (D) TOPOLOGY: | 06 amino acid <i>s</i> no acid | | | | | | | |
| | (ii) HOLECULE TYPE: | protein | | | | | | | |
| 15 | (iii) HYPOTHETICAL: | YES | | | | | | | |
| | (vi) ORIGINAL SOURC (A) ORGANISM: | E: Porphyromonas gingi | valis | | | | | | |
| 20 | (ix) FEATURE: (A) HAME/KEY: (B) LOCATION | misc_feature 1106 | | | | | | | |
| | (xi) SEQUENCE DESCR | IPTION: SEQ ID NO:30 | 2 | | | | | | |
| 25 | Leu Ser Tyr Ser Gly Glu 1 5 | Ser Asp Ala Lys Glu 10 | Ser Asp Gln Asn Cys 15 | | | | | | |
| | Arg Lys Cys Thr Phe Ile | | Val Asn Thr Met Arg | | | | | | |
| 30 | Leu Ile Lys Ala Phe Leu 35 | | Leu Pro Ile Phe Phe 45 | | | | | | |
| 00 | Tyr Lys Arg Phe Ile Ser | | | | | | | | |
| | Pro Ser Cys Ser Ser Tyr 65 70 | | Arg Lys Tyr Gly Pro | | | | | | |
| 35 | Gly Lys Gly Leu Leu Leu 85 | Ser Ile Lys Arg Ile | Leu Arg Cys His Pro 95 | | | | | | |
| | Trp Gly Gly Ser Gly Tyr 100 | Asp Pro Val Pro 105 | | | | | | | |
| 40 | (2) INFORMATION FOR SEC | ID 110:303 | | | | | | | |
| 4 5 | (i) SEQUENCE CHARA (A) LENGTH: 8 (B) TYPE: ami (D) TOPOLOGY: | 61 amino acids no acid | | | | | | | |
| | (ii) NOLECULE TYPE: | protein | | | | | | | |
| 50 | (iii) HYPOTHETICAL: (vi) ORIGINAL SOURC (A) ORGANISM: | | ıvalis | | | | | | |
| 55 | (ix) FEATURE: (A) NAME/KEY: (B) LOCATION | misc_feature 1861 | | | | | | | |
| | (xi) SEQUENCE DESCR | RIPTION: SEQ ID NO:30 |)3 | | | | | | |
| 60 | Leu Glu Met Ala Tyr Asp 1 5 | 10 | 15 | | | | | | |
| | Tyr Ser Tyr Gln Glu Ala 20 | 25 | 30 | | | | | | |
| 65 | Asp Met Leu Val Leu Gly 35 | 40 | 45 | | | | | | |
| | Asp Ile Met Arg His Ty | 55 | 60 | | | | | | |
| 5 0 | Ile Glu Leu Glu Ala Ile 65 70 | 75 | 80 | | | | | | |
| 70 | Ser Pro Ile Phe Thr Pro | 90 | 95 | | | | | | |
| | Asp Ile Cys Ala Asp Me | 105 | 110 | | | | | | |
| 75 | Leu Leu Ser Ile Leu As 115 | n Ser Thr Gln Glu Se 120 | r Leu Val Gin Lys ile 125 | | | | | | |

| | Phe | | | Gln | Gly | Ile | | Tyr | Asp | Thr | Ile | | Ser | Asp | Tyr | Phe |
|------------|------------|------------------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|-------------|------------|------------|------------|
| | | 130 Gln | | Asn | Pro | Ser 150 | 135 Glu | Gly | Lys | Ser | Pro 155 | 140 Ser | Glu | Met | Glu | Ile 160 |
| 5 | 145 Leu | Asp | G1 ۲ | туг | | Asp | Asn | Asp | Phe | | | Glu | Glu | Asp | | |
| | Ser | Pro | Pro | Ser 180 | 165 Glγ | Asn | Ser | Gly | Thr 185 | 170 Gly | Gly | Glγ | Ser | Gly 190 | 175 Asp | Ala |
| 10 | Pro | Glu | Gln 195 | | Thr | Glу | G1 y | Gly 200 | _ | Thr | Thr | Thr | Thr 205 | | Arg | Ser |
| 10 | GJλ. | Gl _y 210 | | Thr | Pro | Ala | Leu 215 | | Thr | Phe | Gly | Thr 220 | | Ile | Thr | Ala |
| | Иet 225 | Ala | Ala | Ala | Gly | Lys 230 | Leu | Asp | Pro | Val | Val 235 | Glу | Arg | Glu | Gln | G1u 240 |
| 15 | Ile | Glu | Arg | Val | 11e 245 | Gln | Ile | Leu | Ser | Arg 250 | Arg | Γ?.ε | Lys | Asn | Asn 255 | Pro |
| | Vāl | Leu | Il≘ | Gly 260 | Glu | Pro | Gly | Vàl | G1; 265 | Րλε | Ser | Ala | Ile | Val 270 | Glu | Glу |
| 20 | Leu | Ala | Glu 275 | Arg | Ile | Val | Asn | Arg 280 | Lγε | Vāl | Ser | Arg | 11e 285 | Leu | Phe | Asp |
| | _ | 290 | | | | Leu | 295 | | | | | 300 | | _ | | _ |
| or. | 305 | | | | | Glu 310 | | | | | 315 | | | | | 320 |
| 25 | - | - | | | 325 | Ile | | | | 330 | - | | | | 335 | |
| | | _ | | 340 | | Ala | | = | 345 | | - | | | 350 | | |
| 30 | - | | 355 | | | Arg | _ | 360 | | | - | | 3 65 | | | |
| | | 370 | | | | Lys | 375 | | | | | 380 | | | | |
| 35 | 385 | | | | | Pro 390 | | | | | 395 | | | | | 400 |
| 33 | | | | | 405 | Ile | | | | 410 | | | | | 415 | |
| | - | - | | 420 | | Ala | | _ | 425 | | | | | 430 | - | _ |
| 4 0 | | | 435 | | | Phe Val | | 440 | | _ | | | 445 | | | - |
| | | 450 | | | | Ala | 455 | | | | | 460 | | | | |
| 45 | 465 | | | | | 470 Gln | | | | | 475 | _ | | | _ | 480 |
| | | | | | 485 | Gln | | | | 490 | | | | | 495 | |
| | | | - | 500 | | Ly:s | | | 505 | | | | - | 510 | - | _ |
| 50 | | | 515 | | | Leu | | 520 | | | | | 525 | | | |
| | | 530 | | | | Arg | 535 | | | | | 540 | | | | |
| 55 | 545 | • | | _ | | 550 Asp | | _ | | | 555 | • | | | - | 560 |
| | | | | | 565 | Gly | | | | 570 | | | | | 575 | |
| | | | | 580 | | Pro | | _ | 585 | | _ | _ | | 590 | _ | |
| 60 | | | 595 | | | Leu | | 600 | | | | | 605 | | | |
| | | 610 | | | | Met | 615 | | | | | 620 | | | | |
| 65 | 625 Ala | Pro | Pro | Gly | Tyr | 630 Val | G1;· | Tyr | Glu | Glu | 635 Gly | Gly | Gln | Leu | Thr | 640 Glu |
| | | | | Arg | 645 | Pro | _ | | Val | 650 | _ | _ | | Glu | 655 | |
| 70 | Lys | Ala | | 660 Ala | Asp | Val | Phe | | 665 Leu | Leu | Leu | Gln | | 670 Net | Asp | Glu |
| 70 | Glγ | | 675 Leu | Thr | Asp | Ser | | G1?. 680 | Arg | Arg | Val | | 685 Phe | Lys | Asn | Thr |
| | | 690 Ile | Ile | Ile | Thr | Ser | 695 Asn | Val | Gly | Thr | | 700 Gln | Leu | Lγε | Asp | |
| 75 | 705 Gly | G1 rı | Gly | Ile | Gly | 710 Phe | Arg | Ser | G1 u | Lys | 715 Asp | G1 u | Glu | Ala | Asn | 720 Lys |

```
725
                                                  730
       Glu His Ser Arg Ser Val Ile Gln Lys Ala Leu Asn Lys Thr Phe Ser
740 745 750
       Pro Glu Phe Leu Asn Arg Leu Asp Asp Ile Ile Leu Phe Asp Gln Leu
755 760 765
       Gly Lys Thr Glu Ile Arg Arg Met Val Asp Ile Glu Leu Lys Ala Val
770 775 780
       Leu Ala Arg Ile His Arg Ala Gly Tyr Asp Leu Val Leu Thr Asp Glu
785 790 795 800
       Ala Lys Asp Val Ile Ala Thr Lys Gly Tyr Asp Leu Gln Tyr Gly Ala
805 815
10
       Arg Pro Leu Lys Arg Thr Leu Gln Asn Glu Val Glu Asp Arg Leu Thr 820 825 830
       Asp Leu Ile Leu Ser Gly Gln Ile Glu Lys Gly Gln Thr Leu 835 840 845
15
       Ser Ala Arg Asp Gly Glu Ile Ile Val Gln Glu Gln Ala
                                   855
       (2) INFORMATION FOR SEQ ID NO:304
20
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 456 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
25
            (ii) HOLECULE TYPE: protein
           (iii) HTPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature (B) LOCATION 1...456
35
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:304
       Arg Gly Gly Gln Ile Arg Arg His His Thr Asp Ser Ser Arg Gly Ser
40
                                                  10
       Asp Ser Lys Ala Cys Arg Phe Ser Phe His Ala Glu Thr Ile Gly Phe 20 \\
       Ser Asn His Gln Arg Ala Lys Met Asn Tyr Leu Tyr Ile Leu Ile Thr
45
       Leu Leu Leu Ser Gly Phe Phe Ser Gly Ala Glu Ile Ala Phe Leu Ser 50 55
       Ser Asp Lys Leu Arg Leu Glu Leu Asp Arg Asn Arg Gly Asp Leu Thr 65 70 75 80
       Gl; Arg Ala Leu Asn Leu Leu Tyr Arg His Pro Asp Gln Leu Val Thr
85 90 95
50
       Thr Leu Leu Val Gly Asn Asn Ile Val Leu Val Val Tyr Gly Leu Leu
100 105
       Het Ala Gl; Leu Leu Ala Ala Pro Leu Ala Gln Trp Ile Asp Asn Asp
115 120 125
55
       Ala Het Ile Val Val Leu Gln Ser Val Leu Ser Thr Ile Ile Leu
130 135 140
       Phe Thr Gly Glu Phe Leu Pro Lys Ala Ile Phe Lys Thr Asn Ala Asn 145 150 160
       Het Het Het Arg Val Phe Ala Leu Pro Ile Val Ala Ile Tyr Tyr Leu
165 170 175
Leu Tyr Pro Leu Ser Lys Leu Phe Thr Gly Leu Ser Arg Ser Phe Ile
180 185 190
60
       Arg Leu Val Asp Lys Asn Tyr Val Pro Thr Thr Val Gly Leu Gly Arg
195 200 205
       Val Asp Leu Asp His Tyr Leu Ala Glu Asn Het Ser Gly Glu Asn Glu 210 215 220

Gln Asn Asp Leu Thr Thr Glu Val Lys Ile Ile Gln Asn Ala Leu Asp 225 230 240
65
       Phe Ser Gly Ile Gln Val Arg Asp Cys Met Ile Pro Arg Asn Glu Met 245 250 255
70
                                                250
       Ile Ala Cys Glu Leu Gln Thr Asp Ile Glu Val Leu Lys Thr Thr Phe 260 265 270
       Ile Asp Thr Gly Leu Ser Lys Ile Ile Ile Tyr Arg Gln Asn Ile Asp
                                      280
75
       Asp Val Val Gly Tyr Ile His Ser Ser Glu Het Phe Arg Gly Gln Asp
```

```
Trp Gln Lys Arg Ile Asn Thr Thr Val Phe Val Pro Glu Ser Net Tyr 305 310 315 320
       Ala Asn Lys Leu Met Arg Leu Leu Met Gln Arg Lys Lys Ser Ile Ala
325 330 335
       335

Ile Val Ile Asp Glu Leu Gly Gly Thr Ala Gly Met Val Thr Leu Glu 340

Asp Leu Val Glu Glu Ile Phe Gly Asp Ile Glu Asp Glu His Asp Thr 355

360

360

365
10
       Arg Lys Ile Ile Ala Lys Gln Leu Gly Pro His Thr Tyr Leu Val Ser 370 380
       Gly Arg Met Glu Ile Asp Asp Val Asn Glu Arg Phe Gly Leu Ser Leu 385 390 395 400
       Pro Glu Ser Asp Asp Tyr Leu Thr Val Ala Gly Phe Ile Leu Asn Ser
405 415
15
       His Gln Asn Ile Pro Gln Ala Asn Glu Val Val Glu Ile Ala Pro Tyr 420 \hspace{1.5cm} 425 \hspace{1.5cm} 430
       The Phe Thr Ile Leu Arg Ser Ser Ser The Lys Ile Glu Leu Val Lys 435 440 445
20
       Het Ser Ile Asp Asp Gln Ser Asn
        (2) INFORMATION FOR SEQ ID NO: 305
25
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 299 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
30
             (ii) NOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
35
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...2\overline{9}9
40
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:305
       Leu Lys Lys Glu Ile Thr Met Lys Gln Asn Tyr Phe Lys Arg Val Cys 1 	 5 	 10 	 15
       Ser Leu Leu Trp Leu Val Leu Pro Met Leu Ile Met Pro Leu Glu Val 20 25 30
45
       Ala Ala Gl<br/>n Glu Ile Ile Pro As<br/>n Glu Glu Val Leu Glu Ser Leu Thr35 \hspace{1cm} 40 \hspace{1cm} 45
        Phe Val Ala Pro Val Glu Glu Thr Asp Ala Ile Glu Ala Glu Val Glu
50 55 60
50
       Ala Leu Gln Glu Ile Val Ala Thr Glu Glu Ile Ala Glu Gln Ala Val
65 70 75 80
       Arg Ser Tyr Thr Tyr Thr Val Tyr Arg Asp Gly Val Lys Ile Ala Ser 85 \\ 0 \\ 90 \\ 95
55
       Gly Leu Thr Glu Pro Thr Phe Leu Asp Glu Asp Val Pro Ala Gly Glu 100 105 110
        His Thr Tyr Cys Val Glu Val Gln Tyr Gln Gly Gly Val Ser Asp Lys
115 120 125
       Val Cys Val Asp Val Glu Val Lys Asp Phe Lys Pro Val Thr Asn Leu
130 135 140
60
        Thr Gly Thr Ala Ser Asn Asp Glu Val Ser Leu Asp Trp Asp Gly Val
145 150 156
       Glu Glu Lys Ala Glu Glu Pro Ala Ser Asp Lys Ala Val Ser Tyr Asn
165 170 175
65
        Val Tyr Lys Asn Gly Thr Leu Ile Gly Asn Thr Ala Glu Thr His Tyr
180 185 190
        Val Glu Thr Gly Val Ala Asn Gly Thr Tyr Ile Tyr Glu Val Glu Val
195 200 205
        Lys Tyr Pro Asp Gly Val Ser Pro Lys Val Ala Val Thr Val Thr Val 210 215 220
70
        Thr Asn Ser Ser Leu Ser Asn Val Asp Gly Gln Ala Pro Tyr Thr Leu
225 230 230 240
        Arg Val Glu Gly Lys Lys Ile Ile Ala Glu Ala His Gly Met Ile Thr
245 250 255
75
        Leu Tyr Asp Ile Asn Gly Arg Thr Val Ala Val Ala Pro Asn Arg Leu
```

295

```
265
                                                                        270
        Glu Tyr Met Ala Gln Thr Gly Phe Tyr Ala Val Arg Phe Asp Val Gly
                275
                               280
        Asn Lys His His Val Ser Lys Ile Gln Val Arg
  5
        (2) INFORMATION FOR SEQ ID NO:306
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 377 amino acids (B) TYPE: amino acid
10
                    (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
15
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISH: Porphyromonas gingivalis
20
             (ix) FEATURE:
                    (A) HAME/KEY: misc_feature
                    (B) LOCATION 1...377
25
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:306
        Leu Phe Asn His Lys Lys Ser Trp Tyr Glu Thr Phe Gln Phe Tyr Leu 1 5 10 15
        Val Val Phe Arg Ser Gly Phe Ile Leu Leu Cys Glu Asn Thr Leu Ala
20 25 30
30
        Gln Gln Lys Thr Glu Glu Phe Ala Pro Val Ser Asp Leu Arg Ala Glu
35 40
        Ala Tyr Gly Ser Thr Val Phe Leu His Trp Thr Pro Pro Tyr Asp Asn
50 55 60
        Pro Met Ile Pro Leu Ser Glu Ser Phe Glu Ser Gly Ile Pro Ala Ile
65 70 75 80
35
        Trp Lys Thr Ile Asp Ala Asp Gly Asp Gly Tyr Asn Trp Met His Leu
85 90
       Thr Asn Phe Thr Gly Gln Ser Gly Leu Cys Val Ser Ser Ala Ser Tyr
100 105 110

Ile Gly Gly Val Gly Ala Leu Thr Pro Asp Asn Tyr Leu Ile Thr Pro
115 120 125
40
        Glu Leu Lys Leu Pro Thr Asp Ala Leu Val Glu Ile Île Tyr Trp Val
130 135 140
45
        Cys Thr Gln Asp Leu Thr Ala Pro Ser Glu His Tyr Ala Val Tyr Ser
145 150 155 160
        Ser Ser Thr Gly Asn Asn Ala Ala Asp Phe Val Asn Leu Leu Tyr Glu
165 170 175
       Glu Thr Leu Thr Ala Lys Arg Ile Gln Ser Pro Glu Leu Ile Arg Gly
180 185
50
       Asn Arg Thr Gln Gly Val Trp Tyr Gln Arg Lys Val Val Leu Pro Asn
195 200 205
        Asp Thr Lys Tyr Val Ala Phe Arg His Phe Asn Ser Thr Asp Asn Phe
210 215 220
55
       Trp Leu Asn Leu Asp Glu Val Ser Ile Leu Tyr Thr Pro Leu Pro Arg
225 230 235 240
        Arg Ala Pro Cys Pro His Pro Gly Gly Tyr Thr Tyr Ser Val Fhe Arg
245 250 255
       Asp Gly Gln Lys Ile Ala Ser Gly Leu Ser Ala Leu Ala Tyr Ile Asp
260 265 270
60
       Thr Asp Val Pro Tyr Gly Thr Gln 280 Tyr Cys Val Gln Val Asn Tyr 275

Leu Gln Gly Asp Ser Tyr Lys Val Cys Lys Asn Ile Val Val Ala Asn 290 295
       Ser Ala Asn Ile Tyr Gly Ala Asp Lys Pro Phe Ala Leu Thr Val Val 305 310 315 320 320 320 325 335 335 335 335
65
       Asp Ile Arg Gly Arg Leu Ile Ala Ser Gly Cys Asp Thr Leu Arg Tyr 340 345
70
       Lys Ala Glu Asn Gly Phe Tyr Leu Ile Lys Ile Gln Val Asn Gly Thr
355 360 365
       Val Tyr Thr Glu Lys Ile Gln Ile Gln
75
```

WO 99/29870

```
(2) INFORMATION FOR SEQ ID NO: 307
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 849 amino acids
  5
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) HOLECULE TYPE: protein
 10
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
 15
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...849
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:307
20
        Phe Gly Ile Ser Pro Ser Met Lys Lys Ser Phe Leu Leu Ala Ile Val
       Met Leu Phe Gly Ile Ala Met Gln Gly His Ser Ala Pro Val Thr Lys 20 25 30
25
       Glu Arg Ala Leu Ser Leu Ala Arg Leu Ala Leu Arg Gln Val Ser Leu
       Arg Het Gl; Gln Thr Ala Val Ser Asp Lys Ile Ser Ile Asp Tyr Val 50 60
       Tyr Arg Gln Gly Asp Ala Glu Arg Gly Ile Thr Ser Gln Glu Glu Gly 65 70 75
30
       Ser Pro Ala Tyr Phe Tyr Val Ala Asn Arg Gly Asn Asn Glu Gly Tyr
85 90 95
                                              90 -
       Ala Leu Val Ala Ala Asp Asp Arg Ile Pro Thr Ile Leu Ala Tyr Ser
100 105 110
35
       Pro Ile Gly Arg Phe Asp Met Asp Ser Met Pro Asp Asn Leu Arg Met 115 120 125
       Trp Leu Gln Ile Tyr Asp Gln Glu Ile Gly Leu Ile Leu Ser Gly Lys
130 135 140
                                                      140
       Ala Gln Leu Asn Glu Glu Ile Leu Arg Thr Glu Gly Val Pro Ala Glu
145 150 155 160
40
       Val His Ala Leu Met Asp Asn Gly His Phe Ala Asn Asp Pro Met Arg
165 170 175
       Trp Asn Gln Gly Tyr Pro Trp Asn Asn Lys Glu Pro Leu Leu Pro Asn
180 185
45
       Gly Asn His Ala Tyr Thr Gly Cys Val Ala Thr Ala Ala Gln Ile
195 200 205
       Met Arg Tyr His Ser Trp Pro Leu Gln Gly Glu Gly Ser Phe Asp Tyr 210 220
       His Ala Gly Ser Leu Val Gly Asn Trp Ser Gly Thr Phe Gly Glu Met
225 230 235 240
50
       Tyr Asp Trp Ile Asn Met Pro Gly Asn Pro Asp Leu Asp Asn Leu Thr
245 250 255
       Gln Ser Gln Val Asp Ala Tyr Ala Thr Leu Het Arg Asp Val Ser Ala
260 265 270
55
       Ser Val Ser Met Ser Phe Tyr Glu Asn Gly Ser Gly Thr Tyr Ser Val
275 280 285
       Tyr Val Val Gly Ala Leu Arg Asn Asn Phe Arg Tyr Lys Arg Ser Leu
290 295 300
       Gln Leu His Val Arg Ala Leu Tyr Thr Ser Gln Glu Trp His Asp Met
305 310 315 320
60
       Ile Arg Gly Glu Leu Ala Ser Gly Arg Pro Val Tyr Tyr Ala Gly Asn
325 330 335
       Asn Gln Ser Ile Gly His Ala Phe Val Cys Asp Gly Tyr Ala Ser Asp 340 345 350
65
       Gly Thr Phe His Phe Asn Trp Gly Trp Gly Gly Val Ser Asn Gly Phe 355 360 365
       Tyr Lys Leu Thr Leu Leu Ser Pro Thr Ser Leu Gly Ile Gly Glu 370 380
       Gly Ile Gly Phe Thr Ile Tyr Gln Glu Ile Ile Thr Gly Ile Glu Pro
385 390 395 400
70
       Ala Lys Thr Pro Ala Glu Ala Gly Thr Asp Ala Leu Pro Ile Leu Ala
405 415
       Leu Lys Asp Ile Glu Ala Glu Tyr Lys Ser Glu Ser Gly Leu Asn Val
                                         425
75
       Gly Tyr Ser Ile Tyr Asn Thr Gly Glu Glu Gln Ser Asn Leu Asp Leu
```

```
435
                                         440
       Gly Tyr Arg Leu Asn Lys Ala Asp Gly Glu Val Ile Glu Val Lys Thr
450 455 460
       Ser Ser Ile Asn Ile Ser Trp Tyr Gly Tyr Gly Glu His Pro Glu Ser
465 470 475 480
       Phe Ser Leu Ala Pro Asn Gln Leu Ser Gln Gly Ile Asn Thr Ile Thr
485 490 495
       Leu Leu Tyr Arg Arg Thr Gly Thr Glu Gln Trp Glu Pro Val Arg His
500 505 510
       Ala Gln Gly Gly Tyr Val Asn Ser Ile Lys Val Asn Thr Thr Asp Pro
515 520 525
10
       Asn Asn Val Val Val Thr Val Asp Asn Asn Glu Gly Lys Leu Ser Ile
530 540
       Val Pro Asn Ser Phe Val Ala Asp Leu Asn Ser Tyr Glu His Ser Thr
545 550 560
15
       Ile Thr Val Gln Phe Asn Ser Asp Ser Pro Asp Glu Ile Arg Thr Pro
565 570 575
       Val Ala Phe Ala Leu Ser Thr Gly Ala Thr Ala Asp Asp Val Ile Ser
580 585 590
       Leu Gly Trp Val Het Ala Glu Val Pro Gly Gly Ser Ser Asn Tyr Pro 595 600 605
20
       Val Val Trp Ser Lys Asp Val Leu Thr Leu Ser Glu Gly Asp Tyr Thr
610 615 620
       Leu Trp Tyr Arg Phe Ser Ile Asn Asn Gln Lys Asp Glu Trp Lys Lys
625 630 635
25
       Ile Gl; Ser Val Ser Val Lys Thr Pro Thr Glu Tyr Thr His Pro Leu
645 650 655
       Phe Glu Val Gly His Asn Gln Thr Ser Thr Tyr Thr Leu Asp Net Ala 660 665 670
       His Asr Arg Val Leu Pro Asp Phe Thr Leu Lys Asn Leu Gly Leu Pro
675 680 685
30
       Phe Asn Gly Glu Leu Val Val Val Phe Arg Gln Thr Gin Ser Ser
690 695 700
       Gly Ser Leu Trp Ala Ala Gln Glu Thr Val His Ile Lys Gln Gly Glu
705 710 720
35
       Thr Phe Val Tyr Lys Pro Val Val Glu Gly Pro Ile Pro Asp Gly Ser
725 730 735
       Tyr Arg Ala Thr Leu His Ala Phe Val As<br/>n Gly Gl<br/>n Gl<br/>n Gln Leu Tyr 740 \hspace{1.5cm} 745 \hspace{1.5cm} 750
       Leu Lys Gly Lys Arg Asn Tyr Thr Val Lys Ile Val Asn Gly Thr Ala
755 760 765
40
       Val Glu Ala Ile Glu Ser Ser Glu Glu Ile Arg Val Phe Pro Asn Pro
770 775 780
       Ala Arg Asp Tyr Val Glu Ile Ser Ala Pro Cys Ile Pro Gln Glu Thr
785 790 795 800
45
       Ser Ile Ile Leu Phe Asp Leu Ser Gly Lys Ile Val Met Lys Asn Ser
805 810
       Leu Ser Ala Gly His Gly Arg Met Asp Val Ser Arg Leu Pro Asn Gly 820 825 830
50
       Ala Tyr Ile Leu Lys Val Asp Gly Tyr Thr Thr Lys Ile Asn Ile Val
55
       (2) INFORMATION FOR SEQ ID NO:308
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 295 amino acids
                   (B) TYPE: amino acid
60
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
65
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
70
                   (A) MANIE/KEY: misc feature
                   (B) LOCATION 1...295
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:308
```

Thr Lys Lys Thr Leu Met Lys Lys Leu Phe Leu Ser Leu Thr Ser Leu

75

```
Val Met Val Phe Ala Val Ala Ser Cys Asp Ile Ile Asp Lys Asp Gln 20 25 30
       Thr Leu Leu Pro Ala Pro Thr Asn Val Thr Pro Asp Asn Pro Asp Asp 35 40 45
       Asn Pro Ser Glu Ile Asp Ile Thr Gln Thr His Thr Glu Lys Tyr Val
        Leu Ala Glu Glu Phe Thr Gly Gln Lys Cys Leu Asn Cys Pro Lys Gly
65 70 80
 10
       His Arg Lys Leu Ala Ala Leu Lys Glu Gln Tyr Gly Lys Arg Leu Thr
85 90
       Val Val Gl; Ile His Ala Gly Pro Gly Ser Leu Val Pro Pro Leu Phe 100 105 110
       15
       Val Tyr Asp Lys Ser Tyr Lys Thr Trp Asp Val Pro Ile Ala Glu Gln
145 150 155 160
20
       Met Glu Gln Lys Ala Lys Ile Asn Ile Phe Ala Val Ala Glu Tyr Thr
165 170
       Asp Thr Gln Lys Ile Lys Val Thr Val Lys Gly Lys Ile Leu Gly 180 180
       Asn Thr Leu Pro Lys Ser Met Val Gln Val Tyr Leu Leu Glu Asp Lys
195 200 205
25
       Leu Ile Ala Pro Gln Val Asp Gly Asn Thr Thr Val Glu Asn Tyr Glu
210 215 220
       His Asn His Val Leu Arg Gly Ala Val Asn Gly Ile Trp Gly Glu Glu
225 230 235 240
30
       Phe Val Asn Leu Lys Asp Tyr Leu Tyr Thr Tyr Ala Val Glu Pro Leu
245 250 255
       Ser Gly Het Ser Phe Val Ala Glu Asn Tyr Ser Ile Val Ala Phe Val
260 265 270
       Tyr Asp Val Gln Thr Phe Glu Val Tyr Asp Val Val His Val Lys Ile
35
       Asn Pro Gln Ser Asp Gly Lys
       (2) INFORMATION FOR SEQ ID NO: 309
40
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 230 amino acids
                  (P) TYPE: amino acid
                  (D) TOPOLOGY: linear
45
           (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
50
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
55
                  (B) LOCATION 1...230
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:309
      Thr Asn Lys Lys Glu Glu Thr Met Lys Lys Ser Ser Val Val Ala Ser
60
       Val Leu Ala Val Ala Leu Val Phe Ala Gly Cys Gly Leu Asn Asn Met
20 25 30
      Ala Lys Gly Gly Leu Ile Gly Ala Gly Val Gly Gly Ala Ile Gly Ala 40
65
      Gly Val Gly Asn Val Ala Gly Asn Thr Ala Val Gly Ala Ile Val Gly 50 55
      Thr Ala Val Gly Gly Ala Ala Gly Ala Leu Ile Gly Lys Lys Het Asp
65 70 75 80
      Lys Gln Lys Lys Glu Leu Glu Ala Ala Val Pro Asp Ala Thr Ile Gln
85 90 95
70
      Thr Val Asn Asp Gly Glu Ala Ile Leu Val Thr Phe Asp Ser Gly Ile 100 105 110
      Leu Phe Ala Thr Asn Ser Ser Thr Leu Ser Pro Asn Ser Arg Thr Ala
115 120 125
75
      Leu Thr Lys Phe Ala Ala Asn Het Asn Lys Asn Pro Asp Thr Asp Ile
```

```
130
                                   135
                                                          140
       Arg Ile Val Gly His Thr Asp Asn Thr Gly Ser Asp Lys Ile Asn Asp
145 150 155 160
       Pro Leu Ser Glu Arg Arg Ala Ala Ser Val Tyr Ser Phe Leu Asn Ser
165 170 175
 5
                                           170
       Gln Gly Val Ser Met Ser Arg Net Ala Ala Glu Gly Arg Gly Ser His
180 185 190
       Glu Pro Val Ala Asp Asn Ser Thr Val Ala Gly Arg Ser Ala Asn Arg
195 200 205
       Arg Val Glu Val Tyr Ile Leu Pro Asn Ala Lys Met Ile Glu Gln Ala
210 215 220
10
       Gln Gln Gly Thr Leu Lys
15
       (2) INFORMATION FOR SEQ ID NO:310
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 342 amino acids
                   (B) TYPE: amino acid
20
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
25
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
30
                   (A) HAME/KEY: misc_feature
                   (B) LOCATION 1...3\overline{42}
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:310
       Asn Arg Asn Arg Asn Met Ser Lys Lys Ser Ile Leu Leu Cys Cys 1 \phantom{00} 5 \phantom{00} 10 \phantom{00} 15
35
       Ser Leu Cys Phe Ile Ser Ala Thr Lys Ala Val Thr Pro Val Arg Asn
20 25 30
                                         25
       Val Arg Asn Ser Gln Val Asn Ser Lys Ala Lys Thr Glu Arg Thr Lys
35 40 45
40
       Pro Ser Asp Ser Val Arg Tyr Ile Ser Asn Met Ile Ala Asp Arg Leu
50 55 60
       Glu Phe Arg Asn Lys Ile Ser Ser Glu Lys Glu Val Arg Lys Ala Glu
65 70 75 80
       Tyr Glu Asn Arg Leu Ala Met Glu Ala Leu Asn Tyr Pro Ala Ile Asp
85 90 95
45
       Leu Tyr Gly Glu Asp Ser Trp Ser Glu Tyr Val Asn Pro Phe Val Gly
100 105 110
       Ala Gly Thr Asp Val Glu Ile Pro Asn Ser Tyr Asp Ile Asp Cys Ser
115 120 125
50
       Ser Phe Val Met Pro Val Glu Asp Lys Gln Val Thr Ser Gln Phe Gly
130 135 140
       Tyr Arg Arg Arg Phe Gly Arg Met His Tyr Gly Ile Asp Leu Ser Val
145 150 150 160
55
       Asn Arg Gly Asp Thr Ile Arg Ala Ala Phe Asp Gly Lys Val Arg Val
165 170 175
       Arg Ser Tyr Glu Ala Arg Gly Tyr Gly Tyr Tyr Ile Val Leu Arg His
180 185 190
       Pro Asn Gly Leu Glu Thr Val Tyr Gly His Net Ser Arg Gln Leu Val
195 200 205
60
       Asp Glu Asn Gln Ile Val Arg Ala Gly Gln Pro Ile Gly Leu Gly Gly 210 215 220
       Ser Thr Gly Arg Ser Thr Gly Pro His Leu His Phe Glu Thr Arg Phe 225 230 235 240
65
       Het Gly Ile Pro Ile Asn Pro Ser Thr Ile Ile Asp Phe Asp Asn Gly 245 250 255
       Val Pro Leu Arg Asp Ile Tyr Thr Phe Lys Arg Gly Ser Asn Ser Arg
260 265 270
       Tyr Ala Lys Ala Ser Lys Thr Ser Ser Arg Tyr Ala Lys Lys Gly Lys 275 280 285
70
         275
       Lys Gly Arg Gln Ala Ser Ser Pro Met Thr Tyr Arg Ile Lys Lys Gly 290 295 300
       Asp Thr Leu Glu Thr Ile Ala Lys Arg His Gly Thr Ser Val Gln Lys
305 310 315 320
75
       Leu Cys Ala Thr Asn Gly Ile Gly Lys Ser Lys Ile Leu Thr Pro Gly
```

| | Lγε | Ala | Leu | Arg 340 | 325 Ile | Lys | | | | 330 | | | | | 335 | |
|----|-----------|------------|-------------|----------------|------------|---------------|------------|-------------|-------------|-----------|---------------------------|------------|------------|------------|-----------|-----------|
| 5 | (2) | INFO | ORI IAC | NOI | FOR | SEQ | ID N | 10:33 | 11 | | | | | | | |
| 10 | | (i) | (E | A) LE 3) TY | | l: 15 amir | 9 an | nino cid | CS: acio | ds | | | | | | |
| 15 | | |) HYE | | | | - | ein | | | | | | | | |
| 10 | | (vi) | ORI (F | | | | | ohyro | omona | as gi | ingiv | /alis | ₹ | | | |
| 20 | | (in) | | 1) 117 | | | | | ature | • | | | | | | |
| | | (xi |) SEÇ | OLEMO | CE DE | SCRI | PTIC |)N: S | SEQ I | ED NO |):311 | L | | | | |
| 25 | Pro 1 | Ser | Γ.'.ε | Thr | Ile 5 | Πle | Lys | Thr | Met | Ala 10 | Lys | Ile | Asn | Phe | Туг 15 | Ala |
| | | Sly | Val | Ser 20 | | Pro | Arg | Ile | Arg 25 | | Arg | Ile | Val | Gly 30 | | Trp |
| 30 | Ile | Ala | Glu 35 | Val | Суз | Ser | Arg | Tyr 40 | Gly | Lys | Ala | Vāl | Gly 45 | Glu | Ile | Ser |
| | Туг | Leu 50 | Phe | Сув | Asp | Asp | Glu 55 | туг | Ile | Leu | $\Gamma \lambda \epsilon$ | Ala 60 | Asn | Gln | Glu | Phe |
| | Leu 65 | Asp | His | Asp | Туг | Туг 70 | Thr | Asp | Ile | Ile | Thr 75 | Phe | Asp | Ser | Cys | Glu 80 |
| 35 | Ala | Asp | Thr | Val | Asn 85 | Gly | Asp | Leu | Leu | 11€ 90 | Ser | Leu | Asp | Thr | Val 95 | Arg |
| | Ser | Asn | Ala | Arg 100 | Ala | Leu | Asp | Leu | Arg 105 | Tyr | Glu | Asp | Glu | Leu 110 | His | Arg |
| 40 | Vâl | Ile | 11e 115 | His | Gly | Ile | Leu | His 120 | Leu | Суя | Glγ | Leu | Lys 125 | Asp | Lys | Ser |
| | r?.e | Lys 130 | | Glu | Ala | Gln | Met 135 | Arg | Ala | Ala | Glu | Glu 140 | Lys | Ala | Leu | Val |
| 45 | 145 | | | | | 150 | | | Glu | Leu | Ser 155 | Leu | Leu | His | Thr | |
| | (2) | | ORMAI | | | | | | | | | | | | | |
| 50 | | (1 | (E | 4) LE 3) Ti | | 1: 39 amir | 95 ar | nino cid | acio | is | | | | | | |
| | | | OH (| | | | _ | ein | | | | | | | | |
| 55 | | |) HYI | | | | | | | | | | | | | |
| | | (vi) |) ORI (7 | | | | | chyro | omona | as gi | ingi | /alis | Ē | | | |
| 60 | | (ix) | | 411 (7 | | | | | sture | . | | | | | | |
| 65 | | (xi |) SE(| QUENC | CE DI | SCR1 | ETIC | DN: 5 | SEQ] | ED HO | 0:313 | 2 | | | | |
| 00 | Lys 1 | Ser | Lys | Thr | Met 5 | Lys | Val | Lys | T)'r | | ilet | Leu | Thr | Leu | | Gly |
| | _ | Ile | Ala | Leu 20 | Asn | Ala | Ser | Ala | Gln 25 | 10 Glu | Asn | Thr | Val | Pro 30 | 15 Ala | Thr |
| 70 | Gly | Gln | Leu 35 | | Ala | Lys | Asn | Val | Ala | Phe | Ala | Arg | Asn 45 | | Ala | Glγ |
| | Ser | Asn 50 | | Phe | Val | Thr | Leu 55 | | Gly | Gly | Val | Ala | | Gln | Phe | Leu |
| 75 | Asn 65 | | Asn | Asn | Asn | Lys 70 | | Leu | l/et | Asp | Arg 75 | | Gly | Ala | Ile | 617 80 |



```
Ser Leu Ser Val Gly Lys Tyr His Ser Pro Phe Phe Ala Thr Arg Leu
                          85
                                               90
       Gin Ile Asn Gly Gly Gln Ala His Thr Phe Leu Gly Lys Asn Gly Glu
                                        105
                    100
 5
       Gln Glu Ile Asn Thr Asn Phe Gly Ala Ala His Phe Asp Phe Met Phe 115 120 125
       Asp Val Val Asn Tyr Phe Ala Pro Tyr Arg Glu Asn Arg Phe Phe His
130 135 140
       Leu Ile Pro Trp Val Gly Val Gly Tyr Gln His Lys Phe Ile Gly Ser
145 150 155 160
10
       Glu Trp Ser Lys Asp Asn Val Glu Ser Leu Thr Ala Asn Val Gly Val
165 170 175
       Het Met Ala Phe Arg Leu Gly Lys Arg Val Asp Phe Val Ile Glu Ala
180 185 190
15
       Gln Ala Ala His Ser Asn Leu Asn Leu Ser Arg Ala Tyr Asn Ala Lys 195 200 205
       Lys Thr Pro Val Phe Glu Asp Pro Ala Gly Arg Tyr Tyr Asn Gly Phe 210 215 220
       Gln Gly Het Ala Thr Ala Gly Leu Asn Phe Arg Leu Gly Ala Val Gly
225 230 235 240
20
       Phe Asn Ala Ile Unk Pro Met Asp Tyr Ala Leu Ile Asn Asp Leu Asn
245 250 255
       Gly Gln Ile Asn Arg Leu Arg Ser Glu Val Glu Glu Leu Ser Lys Arg
260 265 270
       Pro Val Ser Cys Pro Glu Cys Pro Glu Val Thr Pro Val Thr Lys Thr
275 280 285
25
       Glu Asn Ile Leu Thr Glu Lys Ala Val Leu Phe \Lambdarg Phe Asp Ser His 290 295 300
       Val Val Asp Lys Asp Gln Leu Ile Asn Leu Tyr Asp Val Ala Gln Phe 305 310 315
30
       Val Lys Glu Thr Asn Glu Pro Ile Thr Val Val Gly Tyr Ala Asp Pro
325 330 335
       Thr Gly Asn Thr Gln Tyr Asn Glu Lys Leu Ser Glu Arg Arg Ala Lys 340 345 350
       Ala Val Val Asp Val Leu Thr Gly Lys Tyr Gly Val Pro Ser Glu Leu 355 360 365

Ile Ser Val Glu Trp Lys Gly Asp Ser Thr Gln Pro Phe Ser Lys Lys 370 375 380
35
       Ala Trp Asn Arg Val Val Ile Val Arg Ser Lys
40
                             390
       (2) INFORMATION FOR SEQ ID NO:313
             (i) SEQUENCE CHARACTERISTICS:
45
                   (A) LENGTH: 387 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
50
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromenas gingivalis
55
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...387
60
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:313
       Tyr Lys Het Thr Tyr Arg Ile Het Lys Ala Lys Ser Leu Leu Leu Ala
                                                10
       Leu Ala Gly Leu Ala Cys Thr Phe Ser Ala Thr Ala Gln Glu Ala Thr
20 25 30
65
       Thr Gln Asn Lys Ala Gly Met His Thr Ala Phe Gln Arg Asp Lys Ala
               35
                                   40
       Ser Asp His Trp Phe Ile Asp Ile Ala Gly Gly Ala Gly Met Ala Leu 50 60
70
       Ser Gly Trp Asn Asn Asp Val Asp Phe Val Asp Arg Leu Ser Ile Val
65 70 75
       Pro Thr Phe Gly Ile Gly Lys Trp His Glu Pro Tyr Phe Gly Thr Arg 85 \phantom{0}90\phantom{0} 95
       Leu Gln Phe Thr Gly Phe Asp Ile Tyr Gly Phe Pro Gln Gly Ser Lys 100 105 110
75
```



| | Glu / | Arg | Asn | His | Asn | Туг | Phe | Gl _y | Asn | Ala | His | Leu | Asp | Phe | Met | Phe |
|------------|---------------------|------------|-----------------|------------|-----------------------------|--------------|--------------|-----------------|------------|------------|------|--------------------|------------|------------|------------|-----|
| | Asp : | | 115 | | | | | 120 | | | | | 125 | | | |
| 5 | Ile | 130 | | | | | 135 | | | | | 140 | | | | |
| | 145 Asn <i>i</i> | | | | | 120 | | | | | 155 | | | | | 160 |
| 40 | Val A | | | | 165 | | | | | 170 | | | | | 175 | |
| 10 | Phe A | | Ile | TRO | | | | | 185 | | | | | 190 | | |
| | Thr 1 | լչ։ | 195 | | | | | 200 | | | | | 205 | | | |
| 15 | Thr f | 210 | | | | Lys | 215 | | | | | 220 | | | | |
| | 225 Tyr <i>1</i> | | | | Asn | 230 | | | | | 235 | | | | | 240 |
| 20 | Gln V | | | Glu | 245 | | | | Pro | 250 | | | | | 255 | |
| 20 | Glu E | Pro | T'hr | 250 | | | | | 265 | | | | | 270 | | |
| | Tyr F | Phe | 275 | | | | Ala | 280 | | | | | 285 | | | |
| 25 | Val T | 90 | | | | Glu | 295 | | | | | 300 | | | | |
| | 305 Val V | | | | Ala | 310 | | | | | 315 | | | | | 320 |
| 30 | Lys I | eu | Ser | Glu | 325 Arg | Arg | Alā | Lys | Ala | 330 Val | Ala | Lys | llet | Leu | 335 Glu | Lys |
| 50 | Tyr G | 51 y | Val | 340 Ser | Ala | Asp | Arg | Ile | 345 Thr | Ile | Glu | Trp | Γ?.ε | 350 Glչ | Ser | Ser |
| | Glu G | | 355 Ile | Tyr | Glu | Glu | Asn | 360 Ala | Trp | Asn | Arg | | 365 Val | Vāl | Met | Thr |
| 35 | Ala A 385 | | Glu | | | | 375 | | | | | 380 | | | | |
| | (2) I | NFO | RMAT | IOH | FOR | SEQ | ID N | 10:31 | 4 | | | | | | | |
| 40 | | (i) | (A (B |) LE | E CH NGTH PE: POLO | : 19 amir | 95 am | ino id | | s | | | | | | |
| 4 5 | (| ii) | | | E TY | | | | | | | | | | | |
| | | | | | TICA | | | | | | | | | | | |
| 5 0 | | | | | L so | | | | | | | | | | | |
| 50 | | | (A |) OR | GANI | SH: | Porp | hyro | mona | ₃ gi | ngiv | alis | | | | |
| 55 | (| ix) | FEA (A (B |) IIA | : ME/K CATI | EY: ON 1 | misc | _fea 95 | ture | | | | | | | |
| | | | | | E DE | | | | | | | | | | | |
| 60 | Val I 1 | | | | 5 | | | | | 10 | | | | | 1 F. | |
| 60 | Phe V | | | 20 | | | | | 25 | | | | | 3.0 | | |
| | Leu G | | 20 | | | | | 40 | | | | | 4 F. | | | |
| 65 | Val Ti | U | | | | | 55 | | | | | 50 | | | | |
| | Lys Pi 65 | | | | | / 0 | | | | | 7 F. | | | | | 0.0 |
| 70 | Phe II | | | | 85 | | | | | 90 | | | | | le | Leu |
| 70 | Lys Ly | | - | LUU | | | | | 105 | | | | 1 | he I | Leu | |
| | Leu II | T | .15 | | | | | 120 | | | | 1 | つち | | | |
| <i>7</i> 5 | Gln Ly 13 | γε Α 30 | sp 'l | yr ' | rhr I | .eu | Leu (135 | 3ln 8 | Ser i | ?he \ | | Yr <i>F</i> .40 | Ala I | le s | Ser ' | Thr |



```
Ala Ile Gly Phe Thr Leu Ala Met Val Thr Phe Ala Gly Ile Arg Glu
                             150
                                                  155
       Gln Leu Asp Met Thr Asn Leu Pro Lys Ala Het Lys Gly Ile Pro Ser
                        165
                                             170
       Ala Leu Leu Ala Ala Gly Ile Leu Ala Met Ala Phe Met Gly Phe Ser
                    180
       Gly Ile Ala
10
       (2) INFORMATION FOR SEQ ID NO:315
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 876 amino acids
                  (B) TYPE: amino acid
15
                  (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
20
           (vi) ORIGIHAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...876
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:315
       Tyr Arg Ser Tyr Arg Gly Ile Gly Ser Gly Thr His Ser Pro Asn Leu 1 5 15
30
       Lys Asn Arg Leu Lys Arg Ile Gly Ile Arg Ile Pro Asn Arg His Tyr 20 25 30
       Ile His Ile Lys Pro Ile Lys Pro Lys Asn Lys Met Lys Gln Leu Asn 35 40 45
35
       Ile Ile Ser Phe Ile Ile Ala Phe Leu Phe Leu Gly Thr Ser Ala Ser 50 55 60
       Ala Gln Gln Ser Gly Gly Ser Val Thr Gly Thr Val Val Asp Lys Ser
65 70 75 80
       Ser Lys Glu Pro Ile Ala Tyr Val Gln Val Phe Val Lys Gly Thr Thr
85 90 95
40
       Leu Gly Thr Ser Thr Asp Ala Asn Gly Asn Tyr Ser Ile Lys Gly Ile
100 105 110
       Pro Ser Gly Asn Gln Thr Ile Val Ala Arg Leu Met Gly Tyr Ser Thr
115 120 125
45
       Cys Glu Glu Lys Val His Ile Glu Lys Gly Gly Ser Arg His Val Asp
130 135 140
       Leu Tyr Leu Thr Glu Glu Ile Leu Ser Leu Asp Gly Val Val Val Ser
145 150 160
50
       Ala Asn Arg Asn Glu Thr Phe Arg Arg Gln Ala Pro Ser Leu Val Thr
165 170 175
       Val Leu Ser Pro Glu Leu Phe Leu Lys Thr Asn Ser Thr Asn Leu Ser
180 185 190
       Gln Gly Leu Lys Phe Gln Pro Gly Leu Arg Val Glu Asp Asn Cys Gln 195 200 205
55
       Asn Cys Gly Phe Asn Gln Val Arg Ile Asn Gly Leu Glu Gly Ala Tyr
                      215
       Ser Gln Ile Leu Ile Asp Ser His Pro Ile Phe Ser Ser Leu Ala Gly
225 230 235 240
60
       Val Tyr Gl; Leu Glu Gln Met Pro Ala Asn Het Ile Glu Arg Val Glu
245 250 255
       Val Ile Arg Gly Gly Gly Ser Ala Leu Phe Gly Ser Asn Ala Val Gly 260 265 270
       Gly Val Ile Asn Val Ile Thr Lys Glu Pro Leu Arg Asn Ser Ala Glu
275 280 285
65
       Ile Ser His Ser Thr Met Thr Phe Asp His Ala Lys Gly Trp Gly Ser 290 295 300
       Phe Gln Asn Thr Thr Gln Phe Asn Gly Ser Het Leu Thr Glu Asp Arg 305 310 315 320
70
       Lys Ala Gly Val Met Val Phe Gly Gln His Asn Tyr Arg Pro Gly Gln
325 330 335
       Asp Ile Asp Gly Asp Asn Phe Thr Glu Leu Pro Asn Leu Arg Asn Arg
340 345 350
       Ser Leu Gl; Phe Arg Ser Tyr Tyr Lys Thr Gly Leu Tyr Ser Lys Ala
75
                                      360
```

```
Thir Leu Glu Tyr His Ser Het Gln Glu Tyr Arg Arg Gly Gly Asp Arg 370 375 360
       Leu Asp Asn Pro Pro Phe Glu Ala Gln Ile Ala Glu Tyr Leu Gln His
385 390 395
       Tyr Ile Asn Gly Gly Ser Phe Lys Phe Asp Gln Gly Phe Ser Gly Gly 405 410 415
 5
       Lys Asp Phe Phe Ser Leu Tyr Ala Ser Ala Gln Asp Val Gln Arg Arg 420 425 430
       Ser Tyr Tyr Gly Gly Gly Asp Tyr Thr Glu Asn Leu Leu Asn Gly Ala
435 440 445
10
       Val Gln Ser Gly Ser Thr Glu Ser Asp Glu Tyr Asn Asp Ala Phe Thr
450 455 460
       Ala Leu Thr Ser Tyr Gly Thr Thr Lys Gly Phe Asp Leu Gln Gly Gly 465 470 475
15
       Gly Met Tyr Arg His Thr Phe Gly Glu Asn Trp Asp Phe Thr Gly Gly 485 490
       Leu Glu Tyr Ile Tyr Gly Gln Leu Asp Asp Arg Ser Gly Tyr Arg Pro
       Ser Lys Ile Asp Gln Asn Thr Ser Thr Phe Ser Gln Tyr Asp Gln Leu
515 520 525
20
       Glu Tyr Lys Thr Glu Lys Leu Ser Ala Leu Ile Gly Ala Arg Ile Asp 530 540
       Tyr Val Leu Leu Asn Gln Asp Gly Lys Arg Tyr Ile Asp Pro Leu Phe 545 550 555 560
25
       Ile Phe Ser Pro Arg Ala Asn Val Arg Tyr Asn Pro Asn Lys Asn Leu
565 570 575
       Ser Phe Arg Leu Ser Tyr Ser Glu Gly Phe Arg Ala Pro Gln Tyr Phe
580 590 590
       Asp Glu Asp Leu His Val Glu Leu Ala Gly Gly Thr Pro Ile Ser Arg
595 600 605
30
       Val Leu Ser Pro Asn Leu Lys Glu Glu Arg Ser Arg Ser Ile Ser Ala
610 615 620
       Ser Phe Asp Tyr Tyr His Arg Ala Asp Glu Trp Gln Phe Asn Ile Met
625 630 635
35
       Gly Glu Ala Phe Ser Thr Phe Ile Ser Asn Gln Phe Lys Pro Ser Asp
645 650 655
       Lys Val Glu Thr Thr Ser Asp Gly Lys Glu Trp Ile Ile Arg Thr Ile
660 665 670
       Tyr Asn Asp Lys Asp Gly Val Ser Lys Val Tyr Gly Val Asn Leu Glu
675 680 685
40
       Gly Arg Ile Ala Tyr Asn Lys Ser Phe Asp Leu Gln Leu Gly Gly Thr
690 695 700
       Trp Gln Arg Ser Arg Tyr Gly Ser Ile Tyr Thr Ala Val Glu Ala Asp
705 710 715
       Lys Thr Thr Gly Gln Ala Glu Ile Ser Val Lys Asp Tyr Val Arg Thr
725 730 735
45
       Pro Asn Leu Tyr Gly Tyr Phe Val Ala Thr Val Arg Pro Thr Glu His 740 750
       Phe Ala Ile Asn Leu Ser Gly Thr Phe Thr Gly Lys Met Asp Val Val 755 760 765
50
       His Glu Ala Tyr Glu Gly Asp Ile Pro Ala Glu His Ile Ala Pro Asp
770 780
       Gly Ser Phe Asp Phe Glu Met Asn Gly Gln Gln Phe Lys Gly Leu Ala
785 790 795 800
55
       Glu Gly His Ala Lys Leu Val Lys Thr Pro Ala Phe Ala Asp Ile Asp 805 810 815
       Leu Lys Leu Ser His Asp Phe His Leu Ala Ser Thr Met Thr Leu Glu
820 825 830
       Leu Asn Ala Gly Ile Gln Asn Ile Phe Asn Ser Tyr Gln Lys Asp Thr
835 840 845
60
       Asp Lys Gly Pro Gly Arg Ala Ser Thr Tyr Val Tyr Gly Pro Met Gln
850 855 860
       Pro Arg Arg Ile Phe Val Gly Thr Lys Ile Asn Phe
65
       (2) INFORMATION FOR SEQ ID NO:316
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 899 amino acids
70
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
75
           (iii) HYPOTHETICAL: YES
```

```
(vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Perphyromonas gingivalis
            (ix) FEATURE:
                   (A) HAME/KEY: misc feature
                    (B) LOCATION 1...899
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:316
10
       Ile Leu Ash His Leu Arg Lys Thr Met Tyr Lys Lys Ile Ile Ala Val
        Ala Ala Leu Phe Cys Ala Ser Ile Gly Ile Leu Lys Gly Gln Ser Ser
20 25 30
15
        Asp Leu Thr Pro Gln Asp Thr Ile Tyr Ser Pro Glu Ile Ser Tyr Ala
                                         40
        Lys Pro Ile His Lys Thr Ile Ala Ser Ile Glu Ile Glu Gly Met Arg
50 55 60
        Ser Phe Asp Asp Phe Val Leu Arg Asn Leu Ser Gly Leu Ala Val Gly 65 70 70 75
20
        Asp Glu Val Leu Ile Pro Gly Asp Ala Met Ser Ala Ala Val Asn Arg
85 90 95
        Ile Het Arg Gln Gly Tyr Phe Ser Asn Val Arg lie Ile Ala Asp Lys
100 105
        Tyr Val Gly Asn Lys Val Tyr Leu Lys Ile Ile Val Thr Glu Arg Pro
115 120 125
25
        Arg Ile Ser Lys Val Thr Phe Ser Gly Val Lys Lys Ser Glu Arg Glu
130 135 140
        Asp Leu Glu Het Lys Ile Gly Leu Arg Glu Gly Ile Gln Net Thr Arg
145 150 155 160
30
        Asn Asn Glu Asp Lys Val Arg Gln Ile Val Gln Lys Tyr Phe Ser Glu
165 170 175
        Lys Gly Tyr Arg Asp Ala Ser Ile Arg Ile Thr Gln Glu Pro Asp Leu
180 185 190
        Ser Lys Asp Gly Phe Val Asn Val Leu Ile Ser Ile Glu Lys Lys Ser
195 200 205
35
        Lys Thr Lys Val Asn Glu Ile Tyr Phe Ser Gly Asn Lys Ala Leu Ser
210 215 220
        Asn His Lys Leu Arg Met Ala Met Lys Asn Thr Asn Ala Lys Phe Ser
225 230 235 240
40
        Leu Arg Lys His Ile Arg Ser Ser Phe Leu Lys Leu Fhe Ser Thr His 245 250 255
        Lys Phe Val Glu Glu Ser Tyr Arg Glu Asp Leu Val Arg Leu Ile Glu
260 265 270
        Lys Tyr Gln Glu Tyr Gly Tyr Arg Asp Ala Glu IIe Leu Thr Asp Ser 275 280 285

Val Val Lys Ala Pro Asp Gly Lys Arg Val Asp IIe Tyr Leu Asn Ile 290 295 300
45
        Glu Glu Gly Gln Lys Tyr Tyr Ile Lys Asp Val Asn Phe Val Gly Asn 305 310 315 320
50
        Ser Gln Tyr Pro Ser Glu Tyr Leu Glu Arg Val Leu Gly Ile Lys Ser
325
330
335
Gly Asp Val Tyr Asn Gln Arg Arg Leu Ala Lys Arg Leu Asn Glu Asp
340
340
350
 55
        Glu Asp Ala Val Gly Asn Leu Tyr Tyr Asn Asn Gly Tyr Ile Phe Ala 355 360 365
         Trp Val Asp Pro Val Glu Thr Asn Val Val Gly Asp Ser Val Ser Leu
370 375 380
         Asp Ile Arg Ile Ala Glu Gly Lys Gln Ala Asn Ile Asn Lys Val Ile
385 390 395 400
 60
         lle Lys Gly Asn Thr Val Val Tyr Glu Asp Val Val Arg Arg Glu Leu
405 415
         Tyr Thr Lys Pro Gly Gln Leu Phe Ser Arg Glu Asp Ile Ile Asn Ser
420 425 430
         Ile Arg Leu Ile Asn Gln Leu Gly His Phe Asp Ala Glu Lys Ser Ile
435 440 445
 65
         Pro Arg Pro Ile Pro Asn Pro Glu Thr Gly Thr Val Asp Ile Glu Tyr 450 455 460
         Asp Leu Val Pro Arg Ser Ser Asp Gln Leu Glu Leu Ser Val Gly Trp
465 470 475 480
 70
         Ser Gln Ser Gly Leu Leu Phe Arg Gly Ala Ile Lys Phe Thr Asn Phe
485 490 495
         Ser Val Gly Asn Leu Leu His Pro Ser Met Tyr Lys Lys Gly Ile Ile
500 505 510
 75
         Pro Gln Gly Asp Gly Gln Thr Leu Ser Leu Ser Ala Gln Thr Asn Gly
```

```
Lys Tyr Tyr Gin Gln Tyr Ser Val Thr Phe Het Asp Pro Trp Phe Gly 530 535
       Gly Lys Arg Pro Asp Met Phe Ser Phe Ser Ala Phe Tyr Ser Lys Thr
545 550 555 560
       Thr Ala Ile Asp Ser Lys Phe Tyr Ash Ser Ash Ala Gly Ash Tyr Tyr 565 575
       Asn Ala Tyr Tyr Asn Ser Tyr Tyr Asn Asn Tyr Asn Ser Tyr Tyr Asn
580 585 590
10
       Gly Net Ser Asn Tyr Thr Gly Asp Leu Tyr Thr Gln Ala Ser Asp Pro 595 600 605
       Asp Arg Ser Leu Gln Met Leu Gly Thr Ser Ile Gly Tyr Gly Lys Arg
610 615 620
       Leu Thr Trp Pro Asp Asn Trp Phe Gln Ile Tyr Thr Ser Leu Asn Tyr
625 630 635 640
15
       Thr Tyr Tyr Arg Leu Arg Asn Trp Ser Tyr Asn Thr Phe Gln Asn Phe 645 650 655
       His His Gly Ser Ala Asn Asp Leu Asn Leu Glu Leu Arg Leu Ser Arg
660 665 670
20
       Thr Ser Ile Asp Asn Pro Ile Tyr Thr Arg Ser Gly Ser Asp Phe Met 675 680 685
       Val Ser Val Ala Ala Thr Leu Pro Tyr Ser Leu Trp Asp Asn His Asp
690 695 700
       Tyr Ala Ser Gln Asn Leu Ser Val Ser Asp Arg Tyr Arg Phe Ile Glu
705 710 715 720
25
       Tyr His Lys Trp Lys Phe Arg Gly Arg Val Phe Thr Pro Leu Leu Asn 725 730
       Pro Ala Thr His Lys Tyr Thr Pro Val Leu Met Ser Arg Val Glu Gly 740 750
30
       Ala Val Leu Gly Ser Tyr Asn Ser Asn Lys Lys Ser Pro Phe Gly Thr
755 760 765
       Phe Tyr Het Gly Gly Asp Gly Met Ser Ser Tyr Tyr Gly Gly Tyr Met 770 780
       Ash Glu Thr Ile Gly Leu Arg Gly Tyr Lys Ash Gly Ser Ile Ala Gly 785 790 795 800
35
       Asn Asn Tyr Asp Tyr Ala Tyr Ala Tyr Met Arg Leu Thr Met Glu Leu
805 810 815
       Arg Phe Pro Ile Leu Phe Glu Asn Ser Phe Asn Ala Trp Leu Leu Ala
820 825 830
40
       Phe Ala Glu Ala Gly Asn Ala Trp Arg Ser Ile Asp Asn Tyr Asn Pro
835 840
       Phe Asn Leu Lys Arg Ser Ala Gly Val Gly Leu Arg Val Thr Leu Pro
850 855 860
      Met Val Gly Met Leu Gly Ile Asp Trp Gly Tyr Gly Phe Asp Arg Pro 865 870 875
45
       Asp Asn Ser Leu Gln Arg Gly Gly Ser Asn Val His Phe Val Leu Gly
       Gln Glu Phe
50
       (2) INFORMATION FOR SEQ ID NO:317
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 177 amino acids
55
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
60
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
65
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...177
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:317
70
       Gln Ile Asn Met Asn Gly Asp Met Lys Arg Phe Leu Ile Leu Ile Gly
                                              10
                                                                  1.5
       Phe Ala Leu Ala Val Ala Phe Ser Gly Phe Ser Gln Lys Phe Ala Leu
                                        25
75
       Val Asp Met Glu Tyr Ile Leu Arg Asn Ile Pro Asp Tyr Glu Met Met
```

```
40
      Ash Glu Gln Leu Glu Gin Val Ser Lys Lys Trp Gln Ash Glu Ile Glu
50 60
      Ala Leu Glu Asn Glu Ala Gln Ser Het Tyr Lys Lys Tyr Gln Ser Asp
65 76 80
                          70
      Leu Val Phe Leu Ser Ala Ala Gln Lys Lys Thr Gln Glu Glu Ala Ile
95 90 95
      Val Lys Lys Glu Gln Gln Ala Ser Glu Leu Lys Arg Lys Tyr Phe Gly
100 105 110
      Pro Glu Gly Glu Leu Tyr Lys Lys Arg Ser Asp Leu Met Lys Pro Ile
115 120 125
10
      Gln Asp Glu Ile Trp Asn Ala Ile Lys Glu Ile Ala Lys Arg Asn Asn
                             135
      Tyr Gln Net Val Leu Asp Arg Gly Thr Ser Gly Île Île Phe Ala Ser
145 150 155 160
15
      Pro Ser Ile Asp Ile Ser Asp Leu Val Leu Ser Lys Met Gly Phe Ser
20
       (2) INFORMATION FOR SEQ ID NO:318
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 170 amino acids
25
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
30
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
35
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...170
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:318
40
       Arg Ile Asn Lys Gln His Glu Met Lys Lys Phe Phe Leu Met Leu Leu
                                              1.0
       Het Ala Leu Pro Leu Ser Leu Leu Ala Gln Lys Val Ala Val Val Asn
20 25 30
                                        25
                    20
       Thr Glu Glu Ile Ile Ser Lys Met Pro Glu Gln Val Ala Ala Thr Lys
35 40 45
45
       Gln Leu Asn Glu Leu Ala Glu Lys Tyr Arg Leu Asp Leu Lys Ser Met
50 55
       Asp Asp Glu Phe Ala Lys Lys Thr Glu Glu Phe Val Lys Glu Lys Asp 65 70 75
50
       Ser Leu Glu Asn Ile Arg Asn Arg Arg Gln Gln Glu Leu Gln Asp
85 90 95
       Ile Gln Thr Arg Tyr Gln Gln Ser Tyr Gln Thr Met Gln Glu Asp Leu
100 105 110
       Gln Lys Arg Gln Gln Gln Leu Phe Ala Pro Ile Gln Gln Lys Val Ala
115 120 125
55
       Asp Ala Ile Lys Lys Val Gly Asp Glu Glu Asn Cys Ala Tyr Ile Met
130 135 140
       Glu Ala Gly Met Net Leu Tyr Thr Gly Ala Thr Ala Ile Asp Leu Thr
60
                            150
       Ala Lys Val Lys Ala Lys Leu Gly Ile Lys
                        165
       (2) INFORMATION FOR SEQ ID NO: 319
65
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 828 amino acids
                   (B) TYPE: amino acid
(D) TOPOLOGY: linear
70
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
 75
            (vi) ORIGINAL SOURCE:
```

5

240/490

(A) ORGANISM: Porphyromonas gingivalis

```
(ix) FEATURE:
    (A) HAHE/KEY: misc_feature
    (B) LOCATION 1...828
```

(E1) SEQUENCE DESCRIPTION: SEQ ID NO:319

| | | | | | | | | J.,, | O G Q | 10 11 | 0.51 | - | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 10 | Ile 1 | Het | Γ.'.ε | Glu | Ala 5 | Ile | Pro | Arg | Γλε | Asn 10 | ra | T∵r | Ile | r?.e | Leu 15 | Asn |
| | | | Тут | 20 | | | | | 25 | Leu | | | | 30 | Cys | |
| 4- | | | Ala 35 | | | | | 40 | Arg | | | | 4.5 | Val | | |
| 15 | | 50 | Glu | | | | 55 | | | | | 60 | Val | | | |
| | 65 | | Glγ. | | | 70 | | | | | 75 | | | | | 80 |
| 20 | | | Thr | | 85 | | | | | 90 | | | | | 95 | |
| | | | Val | 100 | | | | | 105 | | | | | 110 | | |
| 25 | | | Arg 115 | | | | | 120 | | | | | 125 | | | |
| 23 | | 130 | Thr | | | | 135 | | | | | 140 | | | | |
| | 142 | | Arg | | | 150 | | | | | 155 | | | | | 160 |
| 30 | | | Tyr | | 165 | | | | | 170 | | | | | 175 | |
| | | | Arg | 180 | | | | | 185 | | | | | 190 | | |
| 35 | | | Val 195 | | | | | 200 | | | | | 205 | | | |
| | | 210 | Phe Phe | | | | 215 | | | | | 220 | | | | |
| | 225 | | Tyr | | | 230 | | | | | 235 | | | | | 240 |
| 40 | | | Gln | | 245 | | | | | 250 | | | | | 255 | |
| | | | Thr | 260 | | | | | 265 | | | | | 270 | | |
| 45 | | | 275 Thr | | | | | 280 | | | | | 285 | | | |
| | | 290 | Thr | | | | 295 | | | | | 300 | | | | |
| 5 0 | 305 | | Ser | | | 310 | | | | | 315 | | | | | 320 |
| 50 | | | Glу | | 325 | | | | | 330 | | | | | 335 | |
| | | | Glu | 340 | | | | | 345 | | | | | 350 | | |
| 55 | | Phe | 355 Val | | | | Lys | 360 | | | | | 365 | | | |
| | Phe | 370 Asn | Ser | Asn | Glu | Arg | 375 Glu | Thr | Tyr | Asp | Ile | 380 Gln | Gly | Glu | Tyr | Phe |
| 60 | 385 Leu | Asn | Asp | Val | Gln | 390 Leu | Gly | Ala | Asp | G1 y | 395 Thr | Ala | Ser | Met | | 400 Ser |
| 00 | Glγ | Ser | Glu | Asn 420 | 405 Ser | Asn | Gly | Leu | | 410 Ile | G۱۷ | Arg | Asn | | 415 Glu | His |
| | Ala | Arg | Asn 435 | | Leu | Ser | Туг | Arg 440 | 425 Val | Leu | Asn | Met | | 430 Tyr | Arg | Gly |
| 65 | Glu | Het 450 | Lys | Leu | Asn | Glu | Lys 455 | His | Arg | Leu | Gln | Ala 460 | 445 Gly | Val | Ser | Ala |
| | Gln 465 | Het | Glu | Lγε | Ile | Ala 470 | | His | Ile | | Glu 475 | Trp | Glu | Arg | Arg | |
| 70 | | Vāl | Gly | T7.r | Asn 485 | | Pro | His | Ser | Glu 496 | Thr | Val | Leu | Leu | Met 495 | 480 Tyr |
| | | | Leu | 500 | Alā | | | | 505 | Arg | | | | 510 | Ser | |
| ~- | | | Gln 515 | Asp | | | | 520 | Ser | | | | 575 | Thr | | |
| 7 5 | Leu | 11e | Pro | Gly | lle | Arg | Ala | Ser | Trp | Trp | Ser | Phe | Asn | Lys | Glu | Leu |



```
Leu Val Ser Pro Arg Ile Ser Val Gly Tyr Ser Pro Glu Ser Asn Pro 545 550 560
       Ala Leu Val Leu Arg Ala Ala Ala Gly Leu Tyr Tyr Gln Ala Pro Phe
565 570 575
       Tyr Lys Glu Leu Arg Gln Thr His Lys Asp Ala Glu Gly Asn Asn Val
580 585 590
       Val Val Leu Asn Glu Lys Ile Arg Ser Gln Gly Ala Phe His Ile Leu
595 600 605
10
       Ala Gly Ala Asp Tyr Thr Fhe Glu Met Gly Gly Arg Lys Tyr Lys Phe 610 \, 620 \,
       Thr Ala Glu Ala Tyr Tyr Lys Ser Leu Phe Asn Ile Asn Pro Tyr 11e
625 630 635 640
       lle Glu Asn Val Lys Ile Arg Tyr Leu Gly Glu Asn Ile Gly Ser Gly 645 \hspace{1cm} 650 \hspace{1cm} 655
15
       Tyr Ala Ala Gly Ile Asp Leu Lys Leu Phe Gly Glu Leu Val Pro Gly 660 665 670
       Val Asp Ser Trp Leu Thr Ala Ser Ile Ile Lys Ala Arg Gln Lys Leu
675 680 685
20
       Asp Gly Tyr Gly Ser Leu Pro Leu Met Asn Ala Pro Thr Tyr Asn Phe 690 695 700
       Ser Fhe Phe Leu Gln Glu Tyr Val Pro Gly Asn Lys Arg Ile Thr Ala
705 710 720
       Thr Leu Arg Ala Ala Leu Ser Gly Gly Leu Pro Gln Leu Asn Pro Ser 725 \hspace{1.5cm} 730 \hspace{1.5cm} 735
25
       Lys Gly Leu Ser Ser Pro Ala Phe Thr Ala Pro Ala Tyr Lys Arg Val740 -745 -750
       Asp Leu Gly Val Met Tyr Lys Trp Leu Asp Pro Asp Asp Ser Phe Ala
755 760 765
       Gly Arg Ser Lys Trp Leu Met Gly Val Lys Gly Ala Tyr Ile Gly Ala
770 780
30
       Asp Leu Phe Asn Leu Phe Asp Met Thr Asn Val Asn Ser Tyr Tyr Trp
785 790 795 800
       Val Ser Asp Ala Tyr Gln Gln Gln Tyr Ala Val Pro Asn Tyr Leu Thr
805 810 815
35
       Arg Arg Gln Phe Asn Leu Arg Leu Leu Val Glu Phe
       (2) INFORMATION FOR SEQ ID NO:320
40
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 679 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
45
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
50
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
55
                   (B) LOCATION 1...679
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320
       Pro Ile Pro Phe Leu Leu Asn Met Tyr Ser Gly His His Lys Ile His 1 5 10 15
60
       Tyr Pro Phe Leu Ile Leu Leu Val Cys Leu Ala Phe Ala Ala Cys Lys
20 25 30
       Ser Val Lys Leu Lys Asp Ala Glu Lys Ala His Asp Arg Gln Glu Tyr 35 40 45
65
       Thr Lys Ala Ala Asp Met Tyr Asn Thr Leu Tyr Arg Arg Thr Arg Arg 50 55 60
       Lys Gln Val Glu Met Lys Ala Tyr Thr Ala Phe Arg Ser Gly Glu Asn 65 70 75 80
       Tyr Arg Ala Ala Gly Arg Gln Ala Lys Ala Leu Arg Gly Tyr Leu Asn
85 90 95
70
       Ala Arg Arg Tyr Gly Tyr Pro Asp Ser Val Val Leu Leu Arg Leu Ala
100 105 110
       Gln Thr Tyr Gln Gln Gly Gly Asn Tyr Lys Glu Ala Glu Val Leu Phe
115 120 125
       Arg Gly Tyr Leu Glu Ala Tyr Pro Lys Ser Tyr Phe Ala Ala Ile Gly
```

```
135
       Leu Glu Gly Cys Leu Phe Ala Arg Gln Gln Lys Glu Tyr Pro Thr Arg
145 150 155 160
       Tyr Arg Ile Arg Arg Ala Ala Glu Trp Asn Ser Ala Arg Gly Asp Phe
165 170 175
 5
       Gly Pro Ala Tyr Ala Pro Asp Ala Ser Ala Leu Tyr Phe Thr Ser Ser
180 185 190
       Arg Ser Lys Asp Asp Gly Leu Asp Asn Ser Ser Ile Thr Gly Leu Lys
10
       Pro Asn Asp Ile Tyr Ile Ile Lys Arg Asp Ala Gln Gly Arg Trp Gly 210 215 220
       Arg Pro Asp Ser Val Ser Gly Gly Ile Asn Thr Pro Trp Asp Glu Gly
225 230 235 240
       Val Pro Thr Ile Thr Pro Asp Gly Ser Thr Ile Tyr Tyr Thr Leu Ala
245 255
15
       Gln Gln Gly Ala Asp Tyr Asp Arg Thr Val Gln Ile Tyr Ser Ala Ala
260 265 270
       Arg Ser Gly Glu Gly Gly Trp Ser Asn Gly Ser Leu Val Asp Ile Met 275 280 285
20
       Arg Asp Ser Leu Arg Met Ala Ala His Pro Ser Het Ser Ala Ser Gly 290 295 300
       Asp Tyr Leu Tyr Phe Val Ser Asn Ile Gly Gly Ser Tyr Gly Gly Lys
305 310 320
       Asp Ile Tyr Arg Val Lys Val Ser Asp Arg Ser Tyr Gly Ser Pro Glu
325 330 335
25
       Asn Leu Gly Pro Asp Ile Asn Thr Pro Gly Asp Glu Met Phe Pro Phe 340 350
       Ile Asp Gly Asp Ser Thr Leu Phe Phe Ala Ser Asp Gly His Ala Gly
355 360 365
30
       Leu Gly Gly Leu Asp Ile Phe Lys Ala Thr Leu Asp Ser Thr Gly Gln 370 380
       Trp His Val Val Asn Met Gly Gln Pro Val Asn Ser Ser Ala Asp Asp 395 400
       Phe Gly Leu Ala Val Glu Pro Lys Gly Lys Asn Lys Glu Glu Ala Leu
405 410 415
35
       Pro Asp Asn Gly Val Lys Gly Val Phe Cys Ser Asn Arg Gly Asp Ala
420
430
       Arg Gly Trp Pro His Leu Phe His Phe Glu Leu Pro Ala Ile Tyr Thr
435 440 445
40
       Glu Ile Gin Gly Tyr Val Met Asp Arg Glu Glu Asn Pro Ile Ala Gly
450 455 460
       Ala Thr Val Arg Ile Val Gly Glu Arg Gly Pro Val Gly Gln Gly Phe
465 470 480
       Val Thr Thr Arg Asp Asp Gly Ser Tyr Lys Met Ser Val Gln Gly Asp
485 490 495
45
       Thr Arg Tyr Val Met Leu Ala Gly Ala Ser Gly Tyr Leu Asn Gln Tyr 500 510
       Val Glu Leu Lys Thr Asp Thr Ala Lys Gln Ser Glu Thr Tyr Tyr Val
515 520 525
50
       Asp Phe Phe Leu Ala Ser Arg Glu Lys Ala Glu Gly Leu Gln Asn Ile
530 535 540
       Phe Tyr Asp Phe Asp Lys Ala Thr Leu Arg Pro Glu Ser Met Lys Ser 545 550 550 560
       Leu Asp Glu Leu Ile Arg Ile Leu Thr Asp Asn Pro Asp Ile Arg Ile
565 570 575
55
       Glu Leu Gly Ser His Ala Asp Arg Lys Gly Pro Asp Ala Tyr Asn Leu 580 590
       Gly Leu Ser Asp Arg Arg Ala Lys Ser Val Val Asp Tyr Leu Thr Ser 595 600 605
       Arg Gly Ile Ala Ala Asp Arg Leu Thr Trp Lys Gly Tyr Gly Lys Ser 610 615 620

Val Pro Lys Thr Val Thr Ala Lys Ile Ala Glu Arg His Asp Phe Leu 625 630 635 640
60
       Lys Glu Gly Asp Val Leu Thr Glu Glu Phe Val Ala Pro Leu Thr Glu 645 650 655
65
       Glu Gln Gln Ser Val Cys Asp Gln Leu Asn Arg Arg Thr Glu Phe Arg
                    660
       Val Ile Glu Glu Glu Leu Arg
70
       (2) INFORMATION FOR SEQ ID NO: 321
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 772 amino acids
75
```

(B) TYPE: amino acid

```
(D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
 5
          (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Perphyromenas gingivalis
10
            (ix) FEATURE:
                   (A) NAME/KET: misc_feature
                   (B) LOCATION 1...772
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321
15
       Pro Ala Gln Pro Lys Thr Tyr Cys Ile Arg Tyr Phe Arg Arg Glu Val
1 5 10
       Ser Pro Arg Arg Gln Thr Glu Arg Thr Leu Tyr Ala Asp Lys Met Arg
20 25 30
20
       Arg His Ile Val Lys Cys Cys Pro Ser Ala Ser Ile Leu Leu Gln Glu
       Tyr Ser Thr Thr Glu Leu Phe Ile His His Phe Asn Ile Ser Ile Ile 50 60
       Het Lys Lys Phe Phe Phe Ala Leu Leu Ser Ile Gly Ile Ser Ala Gln 65 70 75 80
25
       Ala Phe Ala Lys Thr Asp Asn Val Pro Thr Asp Ser Leu Arg Val His
85 90 95
       Asn Leu Gln Thr Val Thr Val Tyr Ser Thr Arg Thr Ala Val Pro Leu
100 105
       Lys Lys Ile Pro Ala Lys Met Glu Leu Ile Ser Ser Arg Asn Ile Lys
115 120 125
30
       Gln Ser Gly Phe Asn Asn Met Thr Asp Ile Leu Lys Thr Gln Ser Ser
130 135
       Leu Asp Val Ile Gln Tyr Pro Gly Phe Ser Ser Asn Ile Gly Ile Arg
145 150 155 160
Gly Phe Lys Pro Ser Gly Lys Tyr Val Thr Val Leu Val Asn Gly Ile
165 170 175
35
       Pro Ala Gly Thr Asp Asn Ile Ser Thr Leu Asn Thr Ser Asn Ile Glu
180 185 190
40
       Gln Ile Glu Ile Leu Lys Gly Pro Phe Ser Ser Ile Tyr Gly Thr Asn 195 200 205
       Ala Het Gly Gly Val Val Asn Ile Ile Thr His Lys Ser Lys Asp Lys 210 215 220
       lle His Gly Asn Val Ser Leu Phe Gly Gly Ser Tyr Gln Thr Met Ala
225 230 240
45
       Gly Ser Phe Asn Leu Gly Gly Arg Phe Glu Asp Ile Phe Ser Phe Asp 245 250 255
       Leu Ser Leu Gly Leu Asp Lys Gln Asn Lys Asp Tyr Lys Thr Gly Ser
260 265 270
50
       Asn Asn Phe Leu Ser Leu Ser Lys Leu Glu Glu Ala Ile Val Asp Val 275 280 285
       Asn Ala Thr Lys Asn Lys Lys Met Lys Gly Ser Asp Tyr Thr Val Ala
290 295 300
       Thr Gly Arg Leu Arg Phe Gly Ile Asp Phe Thr Pro Glu Trp Ser Leu 305 310 315 320
55
       Asn Leu Tyr Gln Asn Val Phe Leu Gly Asp Ala Ile Pro Val Gly Gly
325 330 335
       Ser Ile Trp Gly Val Tyr Gly Glu Ser Lys Lys Asn Leu Asn Arg Ser
340 345 350
       Ser Thr Ser Phe Glu Leu Leu Gly Lys His Gly Cys His Thr Leu Gln 355 360 365
60
       Phe Ser Pro Tyr Phe Asn Ile Glu Lys Ser Glu Asn Tyr Asn Asn Ala
370 375 380
       Asp Pro Thr Gly Phe Ile Asn Tyr Lys Ser Asp Tyr Tyr Thr Tyr Gly 385 390 395 400
65
       Ala Leu Leu Gln Asp Lys Ile Ser Phe Gly Gly Gln Asn Ile Val Leu
405 410
       Gly Val Asp Ser Arg Asn Met Thr Met Glu Ser Glu Arg Phe Glu Gln 420 425 430
70
       Ala Gly Val Asn Thr Lys Pro Tyr Asn Pro Gly Tyr Ala Thr Asn Asn 435 440 445
       11e Gly Leu Phe Gly Gln Ala Asn Phe Tyr Leu Leu Asn Asp Ala Leu
450 455 460
                                  455
                                                          460
        Ser Ile Ser Ala Gly Ala Arg Ala Asp Phe Het Phe Phe Asp Leu Lys
75
```

| | Ala | Asn | Glu | Tyr | Leu 485 | Asn | Asn | Glu | Ala | 190 155 | Gln | Glu | Thr | His | Asn 495 | Val |
|------------|-------------|------------|------------|------------|------------|------------|---------------|------------|------------|------------|------------|------------|--------------|--------------------------|------------|------------|
| | lle | Asn | Pro | Asn 500 | Val | Gl7. | Ile | rλæ | Tyr 505 | | Phe | Val | Γ <i>).ε</i> | Gl ₃ . 510 | | Thr |
| 5 | Ala | His | Gly 515 | Thr | Phe | Gly | Ser | Ala 520 | Phe | Ser | Ala | Pro | Asp 525 | | Phe | Gln |
| | Lys | Ala 530 | Glу | Gln | Tyr | Val | G1; 535 | Pro | Phe | G1; | Thr | Thr 540 | | Gly | Asn | Pro |
| 10 | Asp 545 | Leu | L'.ε | Pro | Glu | Lys 550 | Ser | Het | Thr | Trp | Asp 555 | Phe | Gl ? | IJē | G17. | Tyr 560 |
| | | | | | 565 | | Ile | | | 570 | | | | | 575 | |
| | | | | 580 | | | Leu | | 585 | | | | | 590 | | |
| 15 | | | 595 | | | | Asn | 600 | | | | | 605 | | | |
| | Glu | Ala 610 | Leu | Leu | Ser | Туr | Asp 615 | Phe | Gly | Ser | Leu | Phe 620 | Ala | Asn | ГЛE | Phe |
| 20 | 625 | | | | | 630 | Asn | | | | 635 | | | | | 640 |
| | | | | | 645 | | Ala | | | 650 | | | | | 655 | |
| 25 | | | | 660 | | | GJ 7. | | 665 | | | | | 670 | | |
| 25 | | | 675 | | | | Arg | 680 | | | | | 685 | | | |
| | | 690 | | | | | Glu 695 | | | | | 700 | | | | |
| 30 | 705 | | | | | 710 | Leu | | | | 715 | | | | | 720 |
| | | | | | 725 | | Ala | | | 730 | | | | | 735 | |
| 35 | | | | 740 | | | Leu | | 745 | | | | | 750 | | |
| 00 | | | 755 | | туг | HIS | Met | 760 | GIY | Arg | Asn | 11e | 765 | GIÀ | Lys | Val |
| | riec | Val 770 | AEII | rne | | | | | | | | | | | | |
| 4 0 | (2) | INFO | ORI IAT | HOIT | FOR | SEQ | ID I | 10:32 | 22 | | | | | | | |
| | | (i) | | | | | CTERI | | | 46 | | | | | | |
| 4 5 | | | (E | 3) T | PE: | amir | no ac line | cid | 4010 | a c | | | | | | |
| | | (ii) | | | | | prot | | | | | | | | | |
| | | (iii) | | | | | • | | | | | | | | | |
| 50 | | (vi) | ORI | GINA | AL SC | OURCE | Ξ: | | | | | | | | | |
| | | | | | | SM: | Porp | ohyro | omona | as gi | ngiv | /alis | 7 | | | |
| 55 | | (ix) | FEA (A | M (A | VIE/F | KEY: | misc | _fea | ature | 2 | | | | | | |
| | | | | | | | L 4 | | | | | | | | | |
| 60 | <i>-</i> 21 | | | | | | PTIC | | | | | | | | | |
| 00 | 1 | | | | 5 | | Tyr | | | 10 | | | | | 15 | |
| | | | | 20 | | | Met | | 25 | | | | | 30 | | |
| 65 | | | 35 | | | | Leu | 40 | | | | | 45 | | | |
| | | 50 | | | | | Thr 55 | | | | | 60 | | | | |
| 70 | 65 | | | | | 70 | Val | | | | 75 | | | | | 80 |
| . • | | | | | 85 | | Met | | | 90 | | | | | 95 | |
| | | | | 100 | | | Leu | | 105 | | | | | 110 | | |
| 75 | 1)1 | 241 | 115 | inr | Tän | гуѕ | r7.2 | 120 | val | ren | Tyr | ше | 125 | Met | Pro | GLY |

```
The Ser Ser Ser Glu Gly Ile Glu Het Gly Arg Thr His Asn Thr Gln 130\,
       Gly Gly Val Asn Val Ser flet Pro Leu Val Ser Ala Gln Leu Trp Lys
145 150 155 160
       Ser Ile Ala Met Thr Gly Glu Gln Leu Asp Leu Ala Leu Glu Lys Ala
165 170 175
 5
       Arg Ser Ser Arg Ile Asp Leu Val Ala Glu Val Lys Lys Ala Tyr Leu
180 185 190
       Ser Val Leu Leu Ala Glu Asp Ser Tyr Gl\gamma Val Phe Lys Arg Ser Tyr 195 200 205
10
       Asp Asn Ala Leu Ala Asn Tyr Lys Asn Ile Ser Asp Lys Phe Asp Arg 210 215 220
       Gly Leu Val Ala Glu Tyr Asp Lys lie Arg Ala Asn Val Gln Val Arg
225 230 235 240
      Ash Ile Glu Pro Ash Leu Leu Gln Ala Gln Ash Ser Val Ala Leu Ala 245 250 255
Leu Trp Gln Leu Lys Val Leu Met Ser Met Glu Val Glu Thr Pro Ile 260 265 270
15
       Arg Leu Ser Gly Ser Leu Ser Asp Tyr Lys Glu Gln Val Tyr Thr Gly 275 280 285
20
       Tyr Phe Ala Ala Asp Thr Leu Île Ser Asn Asn Ser Ser Leu Arg Gln 290 295 300
       Leu Asp Ile Gln Arg Arg Leu Ala Val Ser Ala Asp Lys Leu Asn Lys 305 310 315 320
25
       Tyr Ser Phe Leu Pro Thr Leu Asn Leu Gly Gly Gln Tyr Thr Tyr Ser 325 330 335
       Pro Phe Ser Thr Ile Ser Leu Ser Leu Tyr Ile Pro Ile Phe Asn Gly 355 360 365
30
       Gly Lys Arg Leu Tyr Asn Val Lys Gln Ser Ala Leu Ser Ile Arg Gln 370 375 380
       lle Asp Leu Gln Arg Arg His Ile Glu Gln Ser Ile Arg Met Gly Ile
385 390 395 400
35
       Lys Asn Gln Asn Asp Arg Leu Arg Thr Cys Met Gln Arg Phe Val Ala
405 410 415
       Ser Glu Glu Ala Val Arg Ser Ala Glu Lys Gly Tyr Gln Ile Ala Glu 420 425 430

Lys Arg Tyr Gln Thr Gly Glu Gly Thr Leu Val Glu Leu Asn Asp Ala 435 440 445
40
       Asp Val Ala Leu Leu Gln Ala Arg Leu Asn Tyr Asn Gln Ala Ile Phe
450 455 460
       Asp Phe Met Thr Ala Lys Ala Glu Leu Asp Lys Met Asn Gly Met Gly
45
       lle Pro Glu Gln
       (2) INFORMATION FOR SEQ ID NO:323
50
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 540 amino acids
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
55
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
60
                    (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...540
65
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:323
        Phe His Asn Phe Asp Phe Leu Asn Gly Ile Lys Leu Phe Ser Met Ala 1 5 10 15
       Asn Asn Thr Leu Leu Ala Lys Thr Arg Arg Tyr Val Cys Leu Val Gly 20 25 30
70
        Phe Cys Trp Leu Met Ala Met Met His Leu Ser Gly Gln Glu Val Thr
                                        40
        Het Trp Gly Asp Ser His Gly Val Ala Pro Asn Gln Val Arg Arg Thr
                                   55
```

| | | 17-3 | | | | _ | | | | | | | | | | |
|------------|-----|------------|-------------|------------|-----------------------------|--------------|------------|-------------|------------|------|------|------------|------------|------------|------|-----|
| | 65 | | | | | 70 | | | Ser | | 7.5 | | | | | 8ú |
| _ | | | | | 85 | | | | Lys | 90 | Thr | | | | Q.F. | Thr |
| 5 | Ala | Leu | Tyr | Leu 100 | Leu | Val | Ser | Asp | Ser 105 | Leu | Ala | Val | Arg | Asp 110 | Leu | Pro |
| | Asp | Туг | ևչ։s 115 | Gly | | Val | Ser | Туг 120 | Asp | Ser | Phe | Pro | 11e 125 | Ser | Lys | Glu |
| 10 | | 130 | | | | | 135 | Ala | Asp | | | 140 | Glγ | | | |
| | 145 | | | | | 150 | | | Pro | | 155 | | | | | 160 |
| 15 | | | | | 165 | | | | Glu | 170 | | | | | 175 | |
| 10 | | | | TRO | | | | | Ala 185 | | | | | 190 | | |
| | | | 195 | | | | | 200 | Asp | | | | 205 | | | |
| 20 | | 210 | | | | | 215 | | Asn | | | 220 | | | | |
| | 225 | | | | | 230 | | | Asp | | 235 | | | | | 240 |
| 25 | | | | | 245 | | | | Gly - | 250 | | | | | 255 | |
| 20 | | | | 260 | | | | | Asn 265 | | | | | 270 | | |
| | | | 275 | | | | | 280 | Leu | | | | 285 | | | |
| 30 | | 290 | | | | | 295 | | Asp | | | 300 | | | | |
| | 303 | | | | | 310 | | | Thr | | 315 | | | | | 320 |
| 35 | | | | | 325 | | | | Сує | 330 | | | | | 335 | |
| 00 | | | | 340 | | | | | Gly 345 | | | | | 350 | | |
| | | | 355 | | | | | 360 | Ile | | | | 365 | | | |
| 4 0 | | 3/0 | | | | | 375 | | Ser | | | 380 | | | | |
| | 202 | | | | | 390 | | | Arg | | 395 | | | | | 400 |
| 45 | | | | | 405 | | | | Met | 410 | | | | | 415 | |
| 43 | | | | 420 | | | | | Ala 425 | | | | | 430 | | |
| | | | 435 | | | | | 440 | Glu | | | | 445 | | | |
| 50 | | 450 | | | | | 455 | | Arg | | | 460 | | | | |
| | 403 | | | | | 4/0 | | | Glу | | 475 | | | | | 480 |
| | | | | | 485 | | | | Trp | 490 | | | | | 495 | Asn |
| 55 | | | | 500 | | | | | Met 505 | | | | | 510 | Gly | |
| | | | 212 | | | | | 520 | Asp | | | | Leu 525 | Val | Phe | Ile |
| 60 | Arg | Phe 530 | Val | Leu | Asp | Asp | Leu 535 | Phe | Asp | Val | Arg | Gln 540 | | | | |
| | (2) | INFC | RHAT | HOI | FOR | SEQ | dI | 10:32 | 24 | | | | | | | |
| 65 | | (i) | (P | () LE | E CH NGTH PE: POLO | : 29 amin | 3 am | nino :id | S: acid | is | | | | | | |
| 70 | | (ii) | 1101 | ECUL | E TY | PE: | prot | ein | | | | | | | | |
| - | (| iii) | HYF | OTHE | TICA | L: Y | ES | | | | | | | | | |
| 7 5 | | (vi) | | | L SO | | | hyrc | mona | s gi | ngiv | alis | | | | |

```
(ix) FEATURE:
                   (A) MAME/KEY: misc feature
                   (B) LOCATION 1...293
 5
           (mi) SEQUENCE DESCRIPTION: SEQ ID NO: 324
      Lys Ser Pro Ser Asp Ser Ala Leu Pro Ser Val Trp Ser Ala Asn Thr 1 \phantom{-}5\phantom{+} 10 \phantom{-}15\phantom{+}
      The Ser Gly Leu Val Gly Gly Lys Arg The Thr Leu Leu Ile Leu Tyr 20 25 30
10
       Ser Net Ala Ile Arg Ser Ser Asn \overline{\mbox{1le}} Glu Ser Ile Gln Cys Phe Val 35 40
      Met Lys Lys Glu Lys Leu Trp Ile Ala Ile Val Ala Gly Leu Ala Phe 50 60
       Val Leu Gly Leu Tyr Ala Leu Gly Arg Ser Val Ala Gln Leu Arg Arg
65 70 80
15
      Ser Gln Pro Ser Val Thr Val Thr Gly Met Ala Glu Arg Asn Phe Lys
85 90 95
Ser Asp Leu Ile Val Trp Thr Ala Ser Tyr Gln Leu Gln Met Met Asp
100 105 110
20
       Leu Glu Ser Ala Tyr Lys Ala Leu Lys Glu Lys Gln Ile Leu Val Ala
115 120 125
       Asp Tyr Leu Lys Asn Lys Gln Leu Pro Asp Ser Ser Tyr Ile Phe Ser
130 135 140
25
       Ser Val Ala Ile Ser Lys Glu Tyr Asn Tyr Tyr Tyr Asp Pro Arg Gln
145 150 150 160
       Glu Gln Asn Val Arg Thr Phe Ala Gly Tyr Leu Leu Ser Gln Thr Val
165 170 175
       Thr Val Thr Ser Gln Asp Ile Glu His Val Glu Lys Ile Ser Arg Asp
180 185 190
30
       Ile Thr Glu Leu Ile Asn Gln Gly Val Glu Ile Thr Ser Asp Arg Pro
195 200 205
       Ala Tyr Tyr Tyr Thr Lys Leu Asn Asp Leu Lys Val Glu Met Leu Arg
210 215 220
35
       Asn Ala Ser Glu Asp Ala Phe Asn Arg Ala Ser Val Ile Ala Glu Gly 225 230 235 240
       Ser Gly Ser Ser Val Gly Lys Met Leu Ser Ser Ser Met Gly Val Phe
245 250 255
       Gln Ile Val Gly Leu Asn Ser Asn Glu Asp Tyr Ser Trp Gly Gly Ser
260 265 270
40
       Phe Asn Thr Ser Ser Lys Met Lys Thr Ala Ser Ile Thr Val Lys Ala
275 280 285
       Ser Phe Ala Leu Lys
45
       (2) INFORMATION FOR SEQ ID NO:325
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 280 amino acids
50
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
55
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
60
            (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...280
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:325
65
        Gly Lys Leu Gln Met Lys Lys Thr Ile Ala Ile Ile Ala Ser Ala Leu
1 5 10
        Leu Ala Leu Gly Ala Val Gly Cys Lys Lys Asn Ala Asp Thr Thr Ala
20 25 30
        Val Ser Glu Lys Asp Ser Ile Ala Leu Ser Het Gly Ile Leu Tyr Gly
35 40 45
        Gln Asp Phe Ala Asn Gln Phe Glu Met Ser Arg Leu Gln Gly Gln Pro
50 55 60
```

Ile Asp Ser Val Ala Phe Leu Asp Gly Phe Lys Tyr Gly Ile Asp Thr 65 70 80

75

```
Thr Arg Phe Ser Tyr Asn Leu Gly Ala Ile Tyr Ala Ser Asn Ile Ala
85 90 95
        Arg Gln Leu Ala His Asp Ser Ile Asp Ile Asp Lys Phe Tyr Ala Ala
100 105 110
       Het Arg Ala Ala Leu Leu Lys Asp Thr Val Ser Ile Ala Met Lys Pro
115 120 125
       Ala Asp Ala Gln Ala Phe Met Gln Arg Ile Gln Ala Lys Lys Gln Arg
130 135 140
       10
       Gly Leu Ala Tyr bys Thr Leu Gln Glu Gly Thr Gly Ala Thr Pro Ser
180 185 190
15
       Leu Ala Asp Thr Val Arg Val Lys Tyr Val Gly Thr Leu Val Asp Gly 195 200 205
       Lys Glu Phe Asp Lys Asn Glu Glu Gly Ile Glu Phe Ala Val Thr Gly 210 215 220
       Val Ile Lys Gly Trp Thr Glu Het Leu Gln Leu Het Lys Val Gly Gln 225 230 235 240
20
       Lys Val Arg Val Val Ile Pro Gln Glu Leu Ala Tyr Gly Glu Thr Gly
245 250 255
       Asn Tyr Thr Ile Glu Pro Phe Ser Thr Leu Thr Phe Glu Met Glu Leu 260 265 270
25
       Ile Gly Ile Lys Pro Gly Lys Lys
       (2) INFORMATION FOR SEQ ID NO:326
30
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 803 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
35
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
40
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...803
45
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326
       Cys Arg Lys Gln Lys Leu Ser Asn Pro Lys Arg Ser Arg Tyr Pro Cys 1 	ilde{5} 5 	ilde{10} 15
50
       Arg Arg Ile Ser Phe Arg Gln Ser Asp L;s Asn Pro Met Lys Val Leu
20 25 30
       Arg Gln Val Phe Leu Pro Ile Leu Phe Val Leu Leu Thr Gly Ala Cys
35 40 45
       Ser Thr Thr Lys Asn Leu Pro Glu Gly Glu Gln Leu Tyr Ile Gly Met 50 55 60
55
       Gly Lys Thr Gln Ile Leu Arg Gln Asp Lys Ser His Ala Gly Gln Gln 65 70 80
       Ala Leu Thr Glu Val Glu Ser Thr Leu Lys Val Thr Pro Asn Gly Ala
85 90 95
60
       Ile Phe Gly Ser Ala Ser Ala Ser Leu Pro Lys Ile Pro Phe Gly Leu 100 105 110
       Trp Leu Tyr Asn Ser Phe Val Gly Asp Ser Thr Val Ile Ser Lys Trp
       Ile Phe Asp Lys Phe Ala Ala Lys Pro Val Phe Ile Ser Gln Val Lys
130 135 140
65
       Ser Asp Ser Arg Ala Lys Val Ala Thr Asn Ile Leu Arg Glu His Gly
145 150 155 160
       Tyr Phe Asp Ala Lys Val Lys Ser Ser Val Thr Thr Leu Lys Lys Asp
165 170 175
70
       Ser Leu Lys Ala Lys Ile Ser Tyr Thr Val Asp Met Ala Ser Pro Tyr
180 185 190
       His Tyr Asp Ser Ile Ile Pro Leu Pro Ile Ser Thr Phe Pro Asp Ser
195 200 205
       Ile Leu Ala Tyr Arg Gln Thr Pro Ser Leu Ile Arg Lys Gly Asp Gln
75
```



| | Dhe | i. ~ | Len | 5.1 s | 1 | I c.u | ui- | (21 s) | c) | D. e. c. | ~1 · | The r | 716 | cor | 5.)- | Len |
|------------|------------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|
| | 225 | U511 | Leo | VIa | ے رب | 230 | піс | Glu | Giu | Mrg | 235 | 1111 | 115 | 2=1 | VIa | 240 |
| | Leu | Arg | Asp | Asn | G1 y 245 | ТУr | T)r | T;r | Phe | Arg 250 | Pro | Gln | Asp | Ile | 11e 255 | T;r |
| 5 | Glu | Ala | Āsþ | Thr 260 | Leu | Leu | Val | Arg | G1; 265 | Alā | Val | C/.a | Leu | Arg 270 | Ala | Lys |
| | L e u | Ser | Glu 275 | Asp | Thr | Pro | Pro | Gln 280 | | Met | Arg | Pro | Trp 285 | Arg | Ile | Gly |
| 10 | re | Arg 290 | Thr | Alā | Val | Leu | Leu 295 | Glу | Met | Asn | G1). | Glu 300 | Ser | Pro | Thr | Asp |
| | Ser 305 | Leu | Glu | Val | Glu | Asp 310 | Net | Lys | Vāl | Leu | Tyr 315 | T'.r | Arg | Lys | Met | Pro 320 |
| | Val | Arg | Pro | Γλε | Ile 325 | Leu | Ala | Гλ.ε | Arg | Phe 330 | Arg | Phe | Phe | Ser | Gly 335 | Asn |
| 15 | Leu | T; r | Arg | Gln 340 | Lys | Asp | Asp | Glu | Thr 345 | Thr | Arg | L','s | Ser | Leu 350 | Ala | Arg |
| | Leu | G1 ? | Ala 355 | Phe | Ser | Val | Ile | Asp 360 | Leu | Asn | Phe | Leu | Gln 365 | Arg | Asp | Ser |
| 20 | Ile | Ser 370 | Gly | Leu | Leu | Asp | Val 375 | Arg | Leu | Leu | Thr | Thr 380 | Leu | Asp | Lys | Pro |
| | Trp 385 | Asp | Ala | Ser | Leu | Glu 390 | Thr | Leu | Phe | Thr | Ser 395 | r?.e | Ser | Asn | Asp | Phe 400 |
| | lle | G1 }. | Fro | Gly | Leu 405 | Asn | Phe | Ala | Leu | Ala 410 | Arg | Arg | Asn | Val | Phe 415 | Gly |
| 25 | GJ?. | G17 | Glu | Asn 420 | Leu | Ser | Trp | Asn | 11e 425 | G17. | Gly | Ser | T7r | Glu 430 | Trp | Glu |
| | Thr | Gly | Asn 435 | Arg | Pro | Glu | Asn | Ser 440 | Ser | Asn | Arg | Leu | 11e 445 | Asp | Ile | Asn |
| 30 | Ser | Tyr 450 | Asn | Met | Asn | Thr | Ala 455 | Val | Asn | Leu | Ser | Phe 460 | Pro | Ser | Ile | Val |
| | Phe 465 | Pro | Gl y | Leu | Leu | Asp 470 | Lys | Tyr | Tyr | Tyr | Tyr 475 | Pro | Thr | Thr | Thr | Thr 480 |
| 0.5 | Phe | Gln | Alā | Ser | Ala 485 | Thr | Ala | Leu | Asn | Arg 490 | Ala | His | Tyr | Phe | Ser 495 | Met |
| 35 | Tyr | Ser | Phe | G1y 500 | Phe | Ser | Thr | Thr | Tyr 505 | Glu | Phe | Gln | Pro | Ser 510 | Lys | Glu |
| | His | Arg | His 515 | Ala | Ile | Phe | Pro | Leu 520 | Ьγε | Leu | Asn | T).r | Asn 525 | Leu | Leu | Gly |
| 40 | | 530 | | | | | 535 | Ala | | | | 540 | | | | |
| | 545 | | | | | 550 | | Phe | | | 555 | | | | | 560 |
| 45 | | | | | 565 | | | Glu | | 570 | | | | | 575 | |
| 40 | | | | 580 | | | | G1 7. | 585 | | | | | 590 | | |
| | | | 595 | | | | | Asp 600 | | | | | 605 | | | |
| 50 | | 610 | | | | | 615 | Thr | | | | 620 | | | | |
| | 625 | Asp | Arg | Asn | GIN | 630 | Leu | Ala | Thr | Arg | Phe 635 | GIY | Thr | GLY | Val | 11e 640 |
| | | | | | 645 | | | Val | | 650 | _ | | | | 655 | |
| 55 | Val | Glγ | Gl y | Ala 660 | Asn | Ser | Ile | Arg | Ala 665 | Phe | Thr | Val | Arg | Ser 670 | Ile | Gly |
| | Pro | Gly | Arg 675 | Phe | Asn | Pro | Asp | Ser 680 | Asp | Asn | Gln | Tyr | Ser 685 | Tyr | Leu | Asp |
| 60 | | 690 | | | | | 695 | | | | | 700 | - | - | _ | Lγε |
| | Leu 705 | Phe | Gly | Asp | Leu | His 710 | Ala | Ala | Val | Phe | Leu 715 | Asp | Alā | Gly | Asn | Val 720 |
| C. | | | | | 725 | | | Ser | | 730 | | | | | 735 | |
| 65 | | | | 740 | | | | Leu | 745 | | | | | 750 | | _ |
| | | | 755 | | | | | 760 | | | | | 765 | | | Val |
| 70 | | 770 | | | | | 775 | | | | | 780 | | | | туг |
| | 785 | | | Arg | Phe | Lys 790 | Asp | Ala | Ile | G17 | Phe 795 | His | Leu | Ala | Vā1 | 800 GJ?. |
| | Tyr | Pro | Phe | | | | | | | | | | | | | |
| <i>7</i> 5 | | | | | | | | | | | | | | | | |

15

250/490

```
(2) INFORMATION FOR SEQ ID NO:327
```

```
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 783 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
 5
                (ii) MOLECULE TYPE: protein
10
              (iii) HYPOTHETICAL: YES
                (vi) ORIGINAL SOURCE:
```

(iz) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION 1...783

| | | | 1. | e) L | OCA1. | ION . | 1 | 183 | | | | | | | | |
|------------|-----|-----|------|------|------------|-------|------|--------|-------|--------|------|-----|-----|-----|-----|-----|
| 20 | | (xi |) SE | QUEN | CE DI | ESCR: | IPTI | 011: 3 | SEQ : | ID III | 0:32 | 7 | | | | |
| | 1 | | | | Tyr 5 | | | | | 10 | | | | | 15 | |
| | | | | 20 | Ile | | | | 25 | | | | | 30 | | |
| 25 | | | 35 | | Arg | | | 40 | | | | | 45 | | | |
| | | 50 | | | Ser | | 55 | | | | | 60 | | | | |
| 30 | 65 | | | | Leu | 70 | | | | | 75 | | | | | 80 |
| | | | | | Arg 85 | | | | | 90 | | | | | 95 | |
| | | | | 100 | Ser | | | | 105 | | | | | 110 | Leu | |
| 35 | | | 115 | | Thr | | | 120 | | | | | 125 | | | |
| | | 130 | | | Phe | | 135 | | | | | 140 | | | | |
| 4 0 | 145 | | | | Lys | 150 | | | | | 155 | | | | | 160 |
| | | | | | Arg 165 | | | | | 170 | | | | | 175 | |
| 45 | | | | 180 | Leu | | | | 185 | | | | | 190 | | |
| 45 | | | 195 | | Ser | | | 200 | | | | | 205 | | | |
| | | 210 | | | Arg | | 215 | | | | | 220 | | | | |
| 50 | 225 | | | | Tyr | 230 | | | | | 235 | | | | | 240 |
| | | | | | Ala 245 | | | | | 250 | | | | | 255 | |
| 55 | | | | 260 | Ile | | | | 265 | | | | | 270 | | |
| 33 | | | 275 | | Phe | | | 280 | | | | | 285 | | | |
| | | 290 | | | Туг | | 295 | | | | | 300 | | | | |
| 60 | 305 | | | | Leu | 310 | | | | | 315 | | | | | 320 |
| | | | | | Asp 325 | | | | | 330 | | | | | 335 | |
| 65 | | | | 340 | Asn | | | | 345 | | | | | 350 | | |
| 00 | | | 355 | | Leu | | | 360 | | | | | 365 | | | |
| | | 3/0 | | | Ala | | 375 | | | | | 380 | | | | |
| 7 0 | 385 | | | | Gly | 390 | | | | | 395 | | | | | 400 |
| | | | | | Phe 405 | | | | | 410 | | | | | 415 | |
| 75 | | | | 420 | Ala | | | | 425 | | | | | 430 | | |
| 75 | Gly | Val | Glu | Ser | Ser | Leu | Arg | Phe | Pro | Arg | Leu | Leu | Phe | Pro | Phe | Ile |

(A) ORGANISM: Porphyromonas gingivalis



| | Ser | Asp 450 | 435 Glu | Thr | Arg | Arg | Arg 455 | 440 Leu | Arg | Ala | Ser | Thr | 445 Glu | Trp | re | Ile |
|------------|------------|-------------|---------------------|------------|---------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|
| 5 | Gly 465 | | Asn | тут | Gln | Thr 470 | | Pro | Glu | Phe | Asp 475 | | Val | Il∈ | Leu | Ser 480 |
| J | | Gln | Leu | Asn | T;r 485 | | Trp | Gln | Thr | Tyr 490 | | His | Aεn | Arg | Leu 495 | |
| | His | Thr | Ile | Arg 500 | | Leu | Asp | Val | Asp 505 | туг | Leu | His | Leu | Pro 510 | туг | lle |
| 10 | Asp | Pro | Asp 515 | Phe | Ala | Gln | Ser | Leu 520 | Pro | Pro | Thr | Thr | Ala 525 | Leu | T).r | Asn |
| | T7.r | Thr 530 | Glu | Gln | Phe | lle | Leu 535 | Gly | Ser | Ala | Tyr | 11e 540 | Leu | Asn | Tyr | Thir |
| 15 | Thr 545 | Ala | Ser | Ser | Het | Glu 550 | Arg | Thr | Val | Ser | Asn 555 | Pro | Phe | Thr | Ala | Arg 560 |
| | Phe | Ser | Ile | Gln | Thr 565 | Alā | Gl y | Asn | Leu | Leu 570 | Gln | Ala | Ile | Ser | Tyr 575 | Leu |
| | Thr | Asp | Ser | Pro 580 | Lys | Asp | Glu | His | G1; 585 | Leu | T;:r | Lys | Met | Phe 590 | G17 | Leu |
| 20 | His | тут | Ala 5 9 5 | Gln | Phe | Val | Lys | Leu 600 | Asp | Leu | Asp | Leu | Ala 605 | Lys | Thr | Vāl |
| | Leu | Leu 610 | Glu | Lys | Asp | Asn | Thr 615 | Leu | Ala | Leu | His | Leu 620 | G17. | Phe | Gly | Leu |
| 25 | Ala 625 | Phe | Pro | Tyr | Gly | Asn 630 | Ala | Arg | His | Ile | Pro 635 | Phe | Glu | Leu | Arg | Tyr 640 |
| | Phe | Ala | Gly | Gl y | Ser 645 | Asn | Ser | Val | Arg | Gly 650 | Trp | Ser | Val | Arg | Thr 655 | Leu |
| 00 | _ | | _ | 660 | | _ | | | Pro 665 | · | = | | | 670 | = | |
| 30 | | - | 675 | | _ | | • | 680 | Asn | | | - | 68Ē | | - | |
| | | 690 | - | | _ | | 695 | | Phe | | - | 700 | _ | | | - |
| 35 | 705 | | _ | | - | 710 | | | Glu | • | 715 | | | - | | 720 |
| | | | | | 725 | | | | Ala | 730 | | | | | 735 | |
| 40 | | | | 740 | | | | | Leu 745 | | | | | 750 | | |
| 40 | | | 755 | | | | | 760 | Lys | | | | 765 | | | Asn |
| | Leu | 770 | Ser | Asn | Phe | Ala | 775 | Hls | Ile | Ala | Val | 780 | Tyr | Pro | Phe | |
| 4 5 | (2) | IHF | ORHA | rion | FOR | SEQ | ID | 110:3 | 28 | | | | | | | |
| | | (i |) SE(| | | | | | CS: acid | 10 | | | | | | |
| 50 | | | (1 | 3) T | YPE: OPOLO | ami | no a | cid | acı | 15 | | | | | | |
| | | <i>t</i> ii |) 1401 | | | | | | | | | | | | | |
| | | | ,) HY! | | | | • | | | | | | | | | |
| 55 | | |) OR: | | | | | | | | | | | | | |
| | | , | | | | | | phyr | omona | as g | ingi | vali | S | | | |
| 60 | | (ix | | A) N | AME/I | | | | atur | e | | | | | | |
| | | | | | OCAT: | | | | | | | • | | | | |
| 65 | V~ 1 | | | | | | | | SEQ : | | | | T1 - | . | C1 = | 7 |
| 03 | 1 | | | | 5 | | _ | | Met Leu | 10 | - | _ | | | 15 | |
| | | | | 20 | | | | | 25 Ile | - | | | | 30 | | |
| 70 | | | 35 | | | _ | • | 40 | Gly | | | | 45 | | | |
| | | 50 | | | | | 55 | | Phe | | _ | 60 | | | _ | _ |
| 75 | 65 | | | | | 70 | | | | _ | 75 | | | | | 80 |
| | 2-1 | val | neu | 1117 | File | vr d | Cys | val | Glγ | ıyr | wig | 1111 | val | 1111 | vià | 361 |



| | Ile | G1γ. | Arg | Gl; | 85 Ser | Gin | Glu | Asp | Leu 105 | 90 Gl; | Thr | Ile | Leu | Leu 110 | 95 Asp | Pro |
|------------|------------|------------|-------------|------------|--------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 5 | Gln | Alā | 11e 115 | Glγ | Leu | Asp | Glu | 11e 120 | | Val | Il€ | Ala | Ser 125 | | Vāl | Pro |
| | Lys | Asp 130 | Arg | Het | Thr | Pro | Val 135 | Pro | Val | Ser | Asn | 11e 140 | Arg | Val | Alā | Asp |
| 40 | 11e 145 | Gln | Alā | Ala | Ser | Leu 150 | Asn | Vāl | Glu | Phe | Pro 155 | Glu | Leu | Val | Γ?.ε | Ser 160 |
| 10 | Thr | Pro | Ser | Thr | T'y'r 165 | Thr | Thr | Lys | Gl y | Ser 170 | G1 7. | Glγ | Phe | Glγ | Asp 175 | Gl ; |
| | | | | Val 180 | | | | | 185 | - | | | _ | 190 | | |
| 15 | | | 195 | Pro | | | | 200 | | | | | 205 | _ | | |
| | | 210 | | Gly | | | 215 | | | | | 220 | | | | |
| | 225 | ren | GIΥ | Ala | ser | ьуs 230 | Leu | GIY | 11e | Ser | 235 | Va1 | GΙΥ | GIY | Thr | Met 240 |
| 20 | | | | Thr | 245 | | | | | 250 | | | | | 255 | |
| | | | | G1γ 260 | | | | | 265 | | | | | 270 | | |
| 25 | Thr | Gly | 11et 275 | Asn | Asp | Gly | Trp | Ala 280 | Ils | Thr | Ile | Ala | Gly 285 | Ser | His | Met |
| | Thr | Gly 290 | Leu | Gly | Туr | Val | Lys 295 | Glу | Leu | L','s | G1?. | Arg 300 | Ala | Phe | Ser | Tyr |
| 00 | 305 | | | Val | | 310 | | | | | 315 | | | | | 320 |
| 30 | Thr | Gly | Phe | G1 y | Ala 325 | Pro | G1n | Trp | His | Asn 330 | Gln | Arg | Ser | Ser | Lys 335 | Tyr |
| | | | | Asp 340 | | | | | 345 | | | | | 350 | | |
| 35 | Gl y | Tyr | Leu 355 | Arg | Gly | Glu | Leu | Thr 360 | Pro | Thr | Ala | Туқ | Ala 365 | Tyr | Asn | Thr |
| | Tyr | His 370 | Lys | Pro | Gln | Phe | Ser 375 | Leu | Asn | His | Phe | Trp 380 | Lys | Met | Asp | Glu |
| 40 | 385 | | | Leu | | 390 | | | | | 395 | | | | _ | 400 |
| 40 | | | | Ala | 405 | | | | | 410 | | | | | 415 | |
| | | | | Gln 420 | | | | | 425 | | | | | 430 | | |
| 45 | | | 435 | Asp | | | | 440 | | | | | 445 | | | - |
| | Ser | 450 | Ala | Ile | Phe | Ala | Leu 455 | Gly | Ser | Asn | Ser | His 460 | Lys | Trp | Phe | Gly |
| F 0 | 465 | | | Ser | | 470 | | | | | 475 | | | | | 480 |
| 50 | | | | Asp | 485 | | | | | 490 | | | | | 495 | |
| | | | | Leu 500 | | | | | 505 | | | | | 510 | | |
| 55 | | | 515 | His | | | | 520 | | | | | 525 | | | |
| | | 530 | | Tyr | | | 535 | | | | | 540 | | | | |
| 60 | 545 | | | Ser | | 550 | | | | | 555 | | | _ | | 560 |
| 60 | | | | Leu | 565 | | | | | 570 | | | | _ | 575 | |
| | | | | Pro 580 | | | | | 585 | | | | | 590 | | |
| 65 | | | 595 | Gly | | | | 600 | | | | | 605 | | | |
| | | 610 | | Gly | | | 615 | | | | | 620 | | | | |
| 70 | 625 | | | Ala | | 630 | | | | | 635 | | | | | 640 |
| 70 | | | | Glu | 645 | | | | | 650 | | | | | 655 | |
| | | | | Asn 660 | | | | | 665 | | | | | 670 | | |
| 75 | Lys | Arg | 11e 675 | Gly | Asn | Glu | Туг | Val 680 | Tyr | Leu | Asn | Glγ | Val 685 | Asp | Ala | Väl |
| | | | | | | | | | | | | | | | | |



```
His Cys Gly Val Glu Ala Glu Val Ser Tyr Arg Pro 11e Arg Gln 11e
690 695 700
       Asp Leu Arg Gly Het Phe Ser Leu Gly Asp Trp Thr Trp Gln Asn Asn 705 710 715 720
       Val Ser Tyr Thr Ser Tyr Asp Glu Ala Gly Asn Glu Thr Gly Gln Asp
725 730 735
 5
       The Thr Tyr The Lys Gly Leu His Val Gly Asp Ala Ala Glm Met Thr 740 745 750
       Ala Ala Val Ser Ala Asp Ile Glu Leu Phe Lys Gly Phe His Val Ile
755 760 765

Gly Lys Tyr Asn Phe Leu Gly Lys Asn Tyr Ala Gly Phe Asn Pro Ala
770 780
10
       Thr Arg Asn Ala Gln Gln Tyr Glu Ala Asp Gly Lys Glu Ile Val Glu
785 790 795 800
       Ser Trp Lys Leu Pro Asp Val Gly Leu Phe Asp Leu Ser Ala Ser Tyr
805 810 815
15
       Asn Phe Lys Leu Gly Ser Leu Ser Thr Thr Phe Tyr Phe Asn Met Asp
820 825 830
       Asn Val Ala Asp Lys Arg Tyr Val Ser Asp Ala Asp Asp Asn Ile Ile
835 840 845
20
       Gly Lys His Asp Glu Ala Ser Ala Leu Val Trp Tyr Gly Phe Gly
       Arg Thr Trp Ser Thr Gly Ile Arg Val Asn Phe
25
       (2) INFORMATION FOR SEQ ID NO: 329
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 460 amino acids (B) TYPE: amino acid
30
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
35
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
40
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                    (B) LOCATION 1...460
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:329
45
       Asn Arg Cys Asn Pro Ser Arg Gln Tyr Phe His Leu Ser Gln Asn Asp
       Val Gly Arg Ser Leu Arg Ala Phe Leu Île His Leu Ser Met Lys Phe
20 25 30
       Ser Ile Arg Leu Phe Leu Cys Ile Ile Phe Leu Leu Ser Ala Phe Ile
35 40 45
50
        Leu Pro Ala Leu Gly Gln Lys Ser Lys Gln Val Gln Arg Leu Glu Lys
50 55 60
       Gln Arg Lys Glu Ala Leu Lys Ala Ile Glu Lys Thr Asp Arg Glu Leu 65 70 80
55
       Arg Asn Thr Lys Lys Asp Lys Gln Asp Lys Gln Lys His Leu Asn Leu
85 90 95
       Leu Asn Lys Gln Val Ala Gln Arg Lys Gln Met Val Gln Leu Leu Asp 100 105 110

Asn Glu Val Lys Glu Leu Gln Ser Asp Ile Asp Ser Met Thr Gly Val 115 120 125
60
        Cys His Gln Leu Ser Val Glu Glu Lys Ala Arg Ser Asp Glu Tyr Ala
130 135 140
        Gln Ala Leu Gln Ser Met Gln Lys Arg Lys Arg Ser Leu Asp Arg Ile
145 150 155 160
65
        Leu Phe Ile Ser Ser Ala Lys Ser Phe Asp Glu Gly Met Arg Arg Met
165 170 175
        Arg Phe Leu Glu Gln Tyr Ala Ser Ala Tyr Lys Leu Ala Ser Val Arg
180 185 190
70
        Leu Arg Asp Thr Arg Ser Lys Leu Glu Thr Glu Arg Ala Thr Val Glu
195 200 205
        Asp Ala Lys Glu Lys Gly His Leu Leu Val Ile Arg Glu Glu Glu
210 215 220
        Lys Lys Leu Glu Gly Gln Gln Ala Glu Gln Arg Arg Gln Val Gln
75
                                230
                                                         235
```

```
Ala Leu Gly Ala Lys Gin Lys Asp Leu Glu Ala Gln Leu Arg Lys Gln 245 250 255
        Lys Lys Gln Ala Glu Ala Leu Asn Arg Lys lie Glu Lys Gln Ile Ala
260 265
        Lys Glu Ile Glu Ala Ala Glu Arg Arg Ala Arg Glu Glu Arg Glu Arg 275 280 285
        Leu Ala Arg Glu Ala Lys Ala Lys Gly Lys Pro Val Pro Ala Glu Pro
290 295 300
       Glu Arg Lys Ala Glu Thr Lys Gly Gly Tyr Ala Het Asp Ala Ser Glu
305 310 315 320
10
       Arg Ala Leu Ser Gly Ser Phe Ala Gln Asn Lys Gly Arg Leu Pro Gly 325 330 335
       Pro Val Arg Gly Arg Tyr Arg Ile Val Ser Asp Phe Gly Val His Gln 340 345 350

His Ser Glu Leu Lys Lys Val Gln Val Asn Asn Gly Gly Ile Asp Ile 355 360 365
15
       Ala Val Ala Thr Gly Ser Asp Ala Thr Ser Val Phe Asp Gly Val Val 370 380
       Ser Ser Val Phe Val Ile Pro Gly Tyr Asn Ser Ala Val Met Val Arg
385 390 395 400
20
       His Gly Asn Tyr Ile Thr Val Tyr Ala Asn Leu Ser Lys Val Tyr Val
405 410 415
       Asn Ser Gly Thr Arg Val Lys Thr Gly Gln Ala Leu Gly Arg Ala Tyr
420 425 430
       Thr Asp Pro Ser Asn Asn Gln Thr Ile 11e His Phe Glu Ile Trp Lys
435 440 445
25
       Glu Arg Ser Lys Gln Asn Pro Arg Leu Trp Leu Arg
30
        (2) INFORMATION FOR SEQ ID NO:330
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 342 amino acids (B) TYPE: amino acid
35
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
40
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
45
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...342
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:330
50
       Ser Phe Tyr Gln Glu Ile Asp Arg Leu Met Lys Lys Tyr Leu Leu Tyr
       Ala Ser Leu Leu Thr Ser Val Leu Leu Phe Ser Cys Ser Lys Asn Asn
20 25 30
       Pro Asn Glu Pro Val Glu Asp Arg Ser Ile Glu Ile Ser Ile Arg Val
55
       Asp Asp Phe Thr Lys Thr Gly Glu Ala Val Arg Tyr Glu Arg Asn Gln
50 60
       Gly Ser Ala Ala Glu Arg Leu Ile Thr Asn Leu Tyr Leu Leu Phe
65 70 80
60
       Asp Gln Ser Gly Ala Asn Pro Ala Lys Tyr Tyr Ile Thr Gly Asn Thr 85 \phantom{\bigg|}90\phantom{\bigg|}95\phantom{\bigg|}
       Phe Thr Gly Gly Thr Trp Leu Pro Asp Asp Met Lys Val Lys Leu Asp 100 105 110
       Net Thr Gln Ser Glu Ala Gly Glu Arg Lys Val Tyr Val Val Ala Asn
115 120 125
65
       Val Asp Asn Ala Val Lys Thr Ala Leu Asp Ala Val Ala Asn Glu Ser
130 135 140
       Asp Leu Gln Thr Val Lys Arg Thr Thr Ala Het Pro Trp Ser Thr Asp 145 150 155 160
70
       Ile Ala Ser Pro Phe Leu Met Ser Gly Asn Lys Thr His Asp Phe Leu
165 170 175
       Ala Asn Arg Leu Leu Asp Asn Val Pro Leu Val Arg Ala Ile Ala Lys
       Val Glu Leu Asn Ile Ser Leu Ser Glu Lys Phe Gln Ile Val Pro Ile
75
                                         200
```

WO 99/29870 PCT/AU98/01023

```
Ile Val Asn Gly Ser Leu Ser Glu Phe Lys Phe Arg Tyr Val Asn Phe
           210
                                  215
                                                       220
       Asp Lys Glu Thr Tyr Val Val Lys Pro Thr Thr Lys Pro Asp Asn Leu
225 230 235 240
       Ile Ser Ser Ala Asn Gly Val Trp Pro Gln Ile Thr Asp Trp Thr Val
245 255
       Trp Gly Ala Ser Leu Asn Thr Ser Pro Ala Pro Asp Ala Gly Thr Gly 260 265 270
       Tyr Thr Leu Asp Ala Asn Gly Lys Val Thr Ala Leu Arg Ile Val Thr
275 280 285
10
       Tyr Leu Asn Glu Arg Asp Ser Lys Gly Ala Thr Val Glu Val Ala Leu
290 295 300
       Pro Arg Val Asp Asp Gly Thr Leu Pro Pro Pro Glu Phe Gly Pro Glu
305 310 315 320
15
       Leu Tyr Arg Leu Pro Leu Pro Asp Lys Ile Leu Arg Asn His Trp Tyr
                     325
                                    330
       Lys Tyr Glu Val Glu Ile
                    340
20
       (2) INFORMATION FOR SEQ ID NO:331
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 329 amino acids
                  (B) TYPE: amino acid
25
                  (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
30
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
35
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...329
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:331
40
       Asn Leu Arg Thr Gln Pro Phe Phe Leu Val Asp Leu Gln Ser Met Ile
                                              10
       Arg Thr Ile Leu Ser Arg Tyr Val Ser Ser Asn Phe Trp Ser Arg Gly
                  20
                                         25
       Ala Thr Phe Phe Phe Thr Ile Phe Pro Ala Phe Ile Leu Ala Ala Thr 35 40 45
45
       Ala Leu Pro Ala Cys Gly Gly Gly Thr Ala Ser Gly Ser Asp Arg Thr
50 55 60
       Leu Ala Val Thr Ile Glu Pro Gln Lys Tyr Phe Ile Glu Ser Ile Ala
65 70 75 80
50
      Asp Lys Ser Val Gln Val Val Ala Leu Val Pro Ala Gly Ser Asn Pro
85 90 95
      Glu Glu Tyr Asp Pro Ser Pro Thr Val Met Lys Arg Leu Ser Glu Ala
100 105 110
      Asp Ala Tyr Phe Tyr Ile Gly Gly Leu Gly Phe Glu Gln Arg Asn Leu
115 120 125
55
      Ala Ala Ile Arg Asp Asn Asn Pro Lys Leu Pro Leu Phe Glu Met Gly
130 135 140
      Lys Ala Leu Ala Asp Ala Gly Ser Ala Asp Leu His Gly Ser Cys Thr
145 150 155 160
60
      Asp His Ser His Thr Asp Leu His Ala His Asp Pro His Tyr Trp Ser
165 170 175
       Ser Val Val Gly Ala Lys Ala Leu Ser Arg Ala Ala Tyr Asp Ala Leu
180 185
      Val Glu Leu Tyr Pro Asn Glu Lys Asp Lys Trp Asp Lys Gly His Asp
195 200 205
65
       Arg Leu Asn Gly Arg Ile Asp Ser Val Lys Arg Leu Val Asp Thr Met 210 215 220
       Phe Ala Asn Gly Lys Ala Asp Lys Ala Phe Val Ile Tyr His Pro Ser 225 230 235 240
70
      Leu Ser Phe Phe Ala Gln Glu Phe Gly Leu Arg Gln Ile Val Ile Glu
245 250 255
      Glu Asp Gly Lys Glu Pro Thr Ala Ala His Leu Arg Arg Val Ile Asp
260 265 270
       Gln Ala Arg Ala Asp Gly Val Arg Ile Val Phe Ile Gln Pro Glu Phe
75
```

PCT/AU98/01023 WO 99/29870

```
Glu Thr Arg Gln Ala Glu Asp Ile Ala Arg Glu Ile Gly Ala Arg Pro 290 -295 -300
       Val Arg Ile Asn Pro Leu Arg Ser Ser Trp Glu Glu Glu Ile Leu His
                          310
 5
       Ile Ala Arg Ala Leu Ala His Glu Arg
                         325
       (2) INFORMATION FOR SEQ ID NO: 332
10
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 878 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
15
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
20
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...878
25
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:332
       Ala Asp Ser Ile Arg Tyr Pro Leu Tyr Phe Phe Gly Arg Asn Arg Lys 1 \phantom{-}5\phantom{+} 10 \phantom{-}15\phantom{+}
30
       Lys Cys Phe Arg Glu Pro Ile Pro Thr Leu Tyr Asn Lys Asn Met Ile
20 25 30
       Gly Lys Lys Ile Phe Phe Ile Leu Leu Ala Leu Ile Ala Phe Ser Gly 35 40
       Leu Asn Ala Ala Thr Asp Thr Glu Phe Lys Tyr Pro Thr Asp Ala Asn 50 60
35
       Ile Ile Gly His Val Lys Asp Ser Lys Thr Gly Glu His Leu Val Gly 65 70 75 80
       Ile Thr Ile Ala Ile Lys Gly Thr Thr Phe Gly Thr Ser Thr Asp Ala 85 90 95
40
       Thr Gly His Tyr Tyr Leu Arg Asn Leu Arg Pro Gly Glu Ile Thr Leu 100 105 110
       45
       Arg Arg Leu Ala Pro Thr Leu Val Asn Val Leu Asn Glu Lys Val Phe 165 170 175
50
       Ser Gln Val Asn Ala Ser Asn Leu Ala Gln Gly Leu Ser Phe Gln Pro
180 185 190
       Gly Val Arg Val Glu Asn Asn Cys Gln Asn Cys Gly Phe Asn Gln Val
195 200 205
       Arg Ile Asn Gly Leu Asp Gly Arg Tyr Ala Gln Ile Leu Ile Asp Ser 210 215 220
55
       Arg Pro Ile Net Ser Ala Leu Ala Gly Val Tyr Gly Leu Glu Gln Ile 225 230 235 240

Pro Ala Asn Met Ile Glu Arg Val Glu Val Val Arg Gly Gly Gly Ser 245 250 250
60
       Ala Leu Tyr Gly Ser Ser Ala Ile Ala Gly Val Val Asn Ile Ile Thr 260 265 270
       Lys Glu Pro Ser His Asn Ser Phe Thr Phe Asn Glu Ser Leu Ser Phe 275 280 285
       Thr Gly Phe Ser Lys Leu Asp Asn Asn Thr Asn Phe Asn Ala Ser Ile
290 295 300
65
       290 295 300
Val Ser Asp Asp Asn Arg Ala Gly Ala Met Val Phe Gly Gln Ala Arg
305 310 315 320
       Tyr Arg Asn His Trp Asp Ala Asn Asn Asp Gly Tyr Ser Glu Leu Gly 325 330 335
70
       Lys Ile Asp Ala Arg Ser Leu Gly Ala His Ser Tyr Leu Arg Leu Ser 340 350
       Asp Tyr Ser Lys Leu Thr Gly Glu Phe His Thr Ile Ser Glu Phe Arg
355 360 365
       Arg Gly Gly Asp Arg Ile Asp Leu Pro Pro His Val Val Gly Val Ala
75
                                  375
```

```
Glu Gln Thr Asp His Ser Val Phe Ser Gly Asn Leu Lys Tyr Asp Leu
385 390 395 400
                                                          395
       Phe Ser Ser Asn Tyr Lys His His Phe Gln Ala Tyr Thr Ser Gly Gln 405 415
       Ile Val Asn Arg Lys Ser Tyr Tyr Gly Gly Ile Gly Glu Ile Asp Val
420 425 430
       Asn Gly His Pro Gly Gly Thr Glu Gly Tyr Pro Ile Pro Gln Asp Gln
435 440 445
       Tyr Gly Asn Asn Tyr Gly Val Thr Lys Gly Lys Thr Tyr Met Gly Gly 450 460

Ile Gln Tyr Ser Tyr Asp Leu Asp Lys Phe Leu Leu Met Pro Ser Gln 465 470 480
10
       Leu Leu Phe Gly Ala Glu Tyr Thr Arg Asp Glu Leu Asn Asp Val Het
405 490 495
       Pro Ile Leu Ser Trp Gln Thr Gly Glu Asp Ala Asn Gly Asn Thr Ile
500 505 510
15
       Pro Leu Tyr Pro Glu Leu Asp Gln Asn Ile Asn Asn Tyr Ser Leu Phe
515 520 525
       Gly Gln Asn Glu Trp Lys Asn Asp Arg Trp Ser Ile Leu Val Gly Ala
530 535 540
20
       Arg Leu Asp Lys His Ser Glu Val Lys Asp Het Ile Leu Ser Pro Arg
545 550 555 560
       Thr Thr Leu Arg Phe Asn Val Asn Pro Asp Ile Asn Leu Arg Ala Thr 565 570 575
       Tyr Ala Lys Gly Phe Arg Ala Pro Gln Val Phe Asp Glu Asp Leu His
580 585 590
25
       Val Gly Val Val Gly Gly Glu Ala Gln Lys Val Phe Asn Asp Pro Asn 595 600 605
       Leu Lys Pro Glu Ile Ser His Ala Phe Ser Leu Ser Ala Asp Het Tyr 610 615 620
30
       His Arg Phe Gly Asn Val Gln Thr Asn Phe Leu Val Glu Gly Phe Tyr 625 630 635 640
       Thr Arg Leu Leu Asp Val Phe Thr Asn Glu Glu Gln Pro Asp Gln His
645 650 655
35
       Asp Gly lie Lys Arg Tyr Thr Arg Ile Asn Gly Ser Gly Ala Lys Val
660 665 670
       Phe Gly Leu Asn Leu Glu Gly Lys Val Ala Tyr Lys Ser Phe Gln Leu
675 680 685
       Gln Ala Gly Leu Thr Leu Ala Ser Asn Lys Tyr Asp Glu Ala Gln Glu
690 695 700
40
       Trp Gly Leu Asn Thr Val Lys Asp Thr Asn Gly Ala Phe Val Thr Glu
705 710 715 720
       Ala Asn Ala Asn Gly Gln Gln Glu Tyr Lys Asn Glu Ser Met Thr Asp
725 730 735
       Thr Gln Ile Thr Arg Thr Pro Ser Val Tyr Gly Tyr Phe Thr Leu Ala 740 750

Tyr Asn Pro Ala His Ser Trp Asn Ile Ala Leu Thr Gly Ala Tyr Thr 755 760 765
       Gly Gln Met Tyr Val Pro His Ala Ile Glu Tyr Gly Val Lys Ser Ala
770 780
50
       Glu Leu Asp Ile Met Gln Asn Asn Pro Glu Ile Thr Asp Glu Thr Gly 785 790 795 800
       Lys Ala Pro Arg Ile Asp Glu Leu Lys Lys Thr Pro Ala Phe Phe Asp
805 810 815
       Leu Gly Leu Lys Val Gly Tyr Asp Phe His Val Phe Gln Ala Thr Glu 820 825 825 830

Val Gln Leu Tyr Val Gly Met Asn Asn Ile Phe Asn Ser Phe Gln Lys 835 840
55
       Asp Phe Asp Arg Gly Ala Ala Arg Asp Ser Gly Tyr Ile Tyr Gly Pro
850 855 860
60
        Thr Gln Pro Arg Thr Gly Tyr Met Gly Leu Val Val Lys Phe
        (2) INFORMATION FOR SEQ ID NO:333
65
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 206 amino acids
                     (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
70
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
75
             (vi) ORIGINAL SOURCE:
```

(A) ORGANISM: Porphyromonas gingivalis (ix) FEATURE: (A) NAME/KEY: misc_feature 5 (B) LOCATION 1...206 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:333 Lys Gln Ile Val Met Thr Val Lys Arg Ala Val Arg Ile Ala Leu Leu 10 10 Thr Leu Ile Gly Ile Leu Phe Ser Ser Pro Ser Leu Val Arg Ala Gln 20 25 30 Ser Leu Phe Ser Thr Glu His Val Leu Gln Leu Tyr Asn Lys Ile Leu 35 40 45 15 Tyr Gly Glu Ser Ala Ala Asp Thr Val Ala Glu Lys Thr Ala Gly Glu 50 60Ser Ala Phe Pro Phe Ile Asp Lys Leu Ile Asn Leu Gly Arg Thr Phe 65 70 75 80Leu Gly Lys Pro Tyr Arg Tyr Arg Gly Pro Ser Pro Trp Pro Met Asp 85 90 95 20 Cys Ser Gly Tyr Val Ser Tyr Leu Tyr Ser Lys Phe Asp Ile Lys Leu 100 105 110 Pro Arg Gly Ala Ala Ala Gln Ser Gln Tyr Thr Asn Pro Ile Glu Arg 115 120 125 Glu Asp Val Arg Pro Gly Asp Leu Leu Phe Phe Lys Gly Arg Asn Ala 130 135 140

Arg Ser Asn Arg Ile Gly His Val Ala Leu Val Val Ser Val Asp Glu 145 150 155 160 25 Asp Asp Ile Thr Met Met His Ser Arg Asn Ser Arg Gly Ile Val Ile 165 170 175 30 Glu Lys Leu Asn Arg Ser Ala Tyr Phe Ser Arg Arg Leu Val Ser Tyr \$180\$ \$190\$Gly Arg Val Pro Gly Ala Lys Arg Val Ile Pro Arg Lys Ser 35 (2) INFORMATION FOR SEQ ID NO: 334 (i) SEQUENCE CHARACTERISTICS: (A) LEHGTH: 467 amino acids
(B) TYPE: amino acid 40 (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein 45 (iii) HYPOTHETICAL: YES (vi) ORIGINAL SOURCE: (A) ORGANISM: Porphyromonas gingivalis 50 (ix) FEATURE: (A) NAME/KEY: misc feature (B) LOCATION $1...4\overline{67}$ (xi) SEQUENCE DESCRIPTION: SEQ ID NO:334 55 Lys Gly Thr Trp Asn Arg Lys Asn Arg Arg Glu Glu Met Lys Arg Thr 1 5 10 15 Ile Leu Leu Thr Ala Leu Thr Val Leu Ser Ser Leu Ser Leu Leu Arg 20 25 30 60 Ala Gln Asn Glu Ser Glu Ala Ser Thr Asn Pro Met Ser Gly Leu Ser 35 40 45 Leu Glu Asp Cys Ile Arg Ile Ala Lys Glu Arg Asn Leu Asn Leu Arg 50 55 Arg Gln Glu Ile Glu Gln Glu Asn Arg Ile Ile Ser Leu Asp Ala Ala 65 70 80 65 Arg His Ser Phe Leu Pro Ser Val Asn Ala Gly Ile Gly His Asn Tyr 85 90 95 Ser Phe Gly Arg Ser Lys Asp Lys Thr Gly Val Thr Val Asp Arg Ser 100 105 110 70 Ser Met Asn Thr Asn Leu Ser Ile Gly Λ la Ser Val Glu Val Phe Ser 115 120 125 Gly Thr Arg Arg Leu His Asp Leu Lys Gln Gln Lys Tyr Asn Val Glu 130 135 140 Asp Gly Ile Ala Arg Leu Gln Lys Ala Arg Glu Asp Leu Ser Leu Gln 75



| | | | | | 165 | | Asn | | | 170 | | | | | 175 | |
|------------|------------|------------|------------|------------|------------|------------|----------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| _ | | | | 180 | | | Ala | | 185 | _ | | | - | 190 | _ | |
| 5 | Ala | Glu | Иet 195 | Val | Arg | Val | Gly | Lys 200 | | Ala | Glu | Gly | Lys 205 | Leu | Leu | Asp |
| | Ile | Asn 210 | Ala | Gln | Иет | Ala | Lys 215 | Asp | Glu | Gln | Leu | Leu 220 | | Gln | Tyr | Arg |
| 10 | Ser 225 | Glu | Glu | Glu | Leu | Ala 230 | Arg | Leu | Asp | Leu | Gly 235 | | Ala | Leu | Glu | Leu 240 |
| | | His | Pro | Glu | Ser 245 | | Ala | Val | Lys | Ala 250 | | Asp | Thr | Asp | Val 255 | Leu |
| | Val | Ala | Glu | Arg 260 | | Gly | Ser | Leu | Leu 265 | | Pro | Glu | Glu | Ile 270 | Tyr | Arg |
| 15 | Thr | Ala | Leu 275 | | Leu | Lys | Pro | Ala 280 | | His | Ser | Ser | Glu 285 | | Gln | Ile |
| | Ala | Ser 290 | Ala | Arg | Glu | Gly | Leu 295 | | Ser | Ala | Arg | Ala 300 | | Tyr | Phe | Pro |
| 20 | Thr 305 | Leu | Ser | Leu | Ser | Ala 310 | Gly | Tyr | Ser | Asn | Gly 315 | Tyr | Phe | Arg | Asp | Leu 320 |
| | | Lys | Glu | Tyr | Ala 325 | | Ile | Asn | Pro | Ser 330 | Phe | Ser | Glu | Gln | | Lys |
| | Asn | Asn | Gly | Ser 340 | | Ser | Ile | Gly | Leu 345 | | Leu | Asn | Ile | | 335 Ile | Phe |
| 25 | Ser | Ala | Met 355 | | Thr | Gln | Asp | Arg 360 | | Arg | Ser | Ser | Arg 365 | 350 Leu | Gln | Ile |
| | Arg | Ser 370 | | Glu | Leu | Arg | Leu 375 | | Glu | Glu | Γλε | Lys 380 | Ala | Leu | Туг | Lys |
| 30 | Glu 385 | | Arg | Gln | Ala | Tyr 390 | Ser | Asn | Ala | Val | Ala 395 | | Asp | Lys | Ala | |
| | | Ala | Ala | Glu | Asn 405 | | Lys | Ala | Ala | Thr 410 | | Lys | Ala | Tyr | | 400 Tyr |
| | Ala | Arg | Asp | Ser 420 | | Glu | Ala | Gly | Arg 425 | | Ser | Ala | Tyr | Glu 430 | 415 Tyr | Ala |
| 35 | Glu | Ala | Lys 435 | | Lys | Tyr | Ala | Leu 440 | | Gln | Val | Glu | Glu 445 | Leu | Arg | Ala |
| | Lys | Tyr 450 | | Phe | Ile | Tyr | Lys 455 | | Lys | Val | Leu | Asp 460 | | Tyr | Gln | Glγ |
| 40 | L√s 465 | | Phe | | | | 100 | | | | | 400 | | | | |
| | | THEC | דבמות | MOTT | FOR | SEO | ID 1 | 10.33 | 15 | | | | | | | |
| | (-, | | | | | | | | | | | | | | | |
| 45 | | (1) | (F | A) LE | NGTH | 1: 45 | CTERI 51 am | nino | | ls | | | | | | |
| | | | (E | | | | no ac line | | | | | | | | | |
| 50 | | (ii) | MOL | ECUI | E TY | PE: | prot | ein | | | | | | | | |
| 30 | (| iii) | HYE | ОТНЕ | TICA | L: Y | ES | | | | | | | | | |
| | | (vi) | | GINA | | | | | | | | | | | | |
| 55 | | | | | | SM: | Porp | hyro | mona | ε gi | ngiv | alis | | | | |
| | | (1X) | (A | | HE/K | | misc | | ture | | | | | | | |
| 60 | | 1 | | | | | 4 | | | | | | | | | |
| 00 | Tura | | | | | | PT10 | | | | | | | | | |
| | Lys 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| 65 | Thr | | | 20 | | | | | 25 | | | | | 30 | | |
| | Asn I | | 35 | | | | | 40 | | | | | 45 | | | |
| 70 | | 50 | | | | | 55 | | | | | 60 | | | | |
| 70 | Leu 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| | Ala | | | | 85 | | | | | 90 | | | | | 95 | |
| 7 5 | Thr | Asp | Ile | Asn 100 | Ser | Phe | Asn | | Ser 105 | Tyr | Ser | Ile . | | Ala 110 | Thr | Het |
| | | | | | | | | | | | | | | | | |

```
Thr Leu Phe Asp Gly Leu Gln Ser Val Tyr Arg Leu Arg Met Ala His 115 120 125
       Ala Arg Arg Glu Ala Ser Arg Leu Ser Val Arg Glu Gln Gln Glu Leu
130 135 140
       Ala Ala Leu Gly Thr Thr Glu Ala Tyr Tyr Asp Leu Val Tyr Ala Arg
145 150 155 160
 5
       Gin Met Gin Glu Leu Ala Met Gin Lys Tyr Glu Glu Ser Ser Arg Leu
165 170 175
       His Arg Gln Thr Ala Arg Met Glu Glu Leu Gly Met Lys Ser Arg Pro
180 185 190
10
       Asp Val Leu Glu Met Gln Ser Arg Met Ala Gl\gamma Asp Arg Leu Ala Leu 195 200 205
       Thr Gln Ala Asp Asn Gln Cys Ile Ile Ala Leu Ile Arg Leu Lys Glu
210 215 220
15
       Lys Met Asn Phe Pro Ile Asp Asp Glu Leu Val Val Asp Asp Met Pro 225 230 235 240
       Ala Asp Ser Leu Ser Ala Asp Met Ala Glu Ser Asp Ser Ser Ala Gly
245 255
       Val Phe Ala Arg Ala Ala His His His Pro Val Leu Leu Arg Ala Lys
260 265 270
20
       Leu Asp Glu Gln Ala Ala Thr Asp Arg Leu Arg Ala Ala Arg Gly Ala
275 280 285
        Phe Leu Pro Ser Val Ser Val Ser Gly Gly Trp Asn Thr Gly Phe Ser 290 295 300
25
       Arg Phe Leu Asn Gly Ser Asp Tyr Thr Pro Phe Ser Glu Gln Phe Arg 305 310 315 320
       Asn Arg Arg Gly Glu Tyr Val Ser Leu Asn Leu Ser Ile Pro Ile Phe
325 330 335
       Ser Gly Phe Ser Leu Val Ser His Leu Arg Gln Ala Arg Ala Glu Arg
340 345 350
30
       Arg Ala Ala Ile Val Arg Arg Gly Glu Ala Glu Arg Arg Leu Tyr Ser
355 360 365
       Glu Ile Ala Gln Ala Met Ala Asp Arg Asp Ala Ala Leu Ala Ser Tyr
370 375 380
       Arg Gln Ala Lys Glu His Thr Asp Ala Met Gln Thr Ala Tyr Glu Ala 385 390 395 400

Val Leu Gln Arg Tyr Glu Glu Gly Leu Asn Thr Ala Ile Asp Leu Thr 405 410 415
35
       Thr Gln Ala Asn Arg Leu Leu Asp Ala Arg Val Gln Arg Leu Arg Ala 420 425 430
40
       Ala Met Thr Tyr Arg Leu Lys Cys Lys Leu Ile Ala Tyr Tyr Gly Cys
               435
        Leu Ser Asp
           450
45
        (2) INFORMATION FOR SEQ ID NO:336
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 962 amino acids
50
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
55
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
60
             (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...962
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:336
65
       Ala Ile Phe Val Val Ser Leu Gln Ile Glu Lys Ile Thr Glu Asn Lys 1 \phantom{\bigg|} 5 \phantom{\bigg|} 10 \phantom{\bigg|} 15
       Tyr Asn Ser Asp Lys Ser Met Asn Lys Phe Tyr Lys Ser Leu Leu Gln 20 25 30
70
       Ser Gly Leu Ala Ala Phe Val Ser Met Ala Thr Ala Leu Thr Ala Ser
35 40 45
       Ala Gln Ile Ser Phe Gly Gly Glu Pro Leu Ser Phe Ser Ser Arg Ser
50 55 60
        Ala Gly Thr His Ser Phe Asp Asp Ala Het Thr Ile Arg Leu Thr Pro 65 70 75 80
75
```



| | | | | | 85 | | | | | 90 | | | | | 95 | Gln |
|------------|-----|------------|-----|-----|-----|-----|------------|-----|-----|-------------|-----|------------|-----|-----|-----|----------------------|
| _ | | | | 100 |) | | | | 105 | , | | | | 110 |) | Val |
| 5 | | | 115 | , | | | | 120 |) | | | | 125 | Asp | Val | Asp |
| | Val | Tyr 130 | Arg | Leu | Gln | Phe | Lys 135 | Leu | Glu | Gly | Ala | Lys 140 | Ala | Ile | Thr | Leu |
| 10 | 145 |) | | | | 150 |) | | | | 155 | Arg | Leu | | | Tyr 160 |
| | | | | | 165 | | | | | Ala 170 | | | | | 175 | His |
| | | | | 180 | | | | | 185 | | | | | 190 | | Leu |
| 15 | | | 195 | | | | | 200 | | Gly | | | 205 | | | |
| | | 210 | | | | | 215 | | | | | 220 | | | | Val |
| 20 | 225 | | | | | 230 | | | | | 235 | | | | | Cys 240 |
| | | | | | 245 | | | | | Ala 250 | | | | | 255 | $L\gamma\varepsilon$ |
| 05 | | | | 260 | | | | | 265 | $L\gamma s$ | | | | 270 | | |
| 25 | | | 275 | | | | | 280 | | Lys | | | 285 | | | |
| | | 290 | | | | | 295 | | | Ile | | 300 | | | | |
| 3 0 | 305 | | | | | 310 | | | | Phe | 315 | | | | | 320 |
| | | | | | 325 | | | | | Ile 330 | | | | | 335 | |
| 25 | | | | 340 | | | | | 345 | Pro | | | | 350 | | |
| 35 | | | 355 | | | | | 360 | | Val | | | 365 | | | |
| | | 3/0 | | | | | 375 | | | Asp | | 380 | | | | |
| 4 0 | 385 | | | | | 390 | | | | Met | 395 | | | | | 400 |
| | | | | | 405 | | | | | Ser 410 | | | | | 415 | |
| ΛE | | | | 420 | | | | | 425 | Asp | | | | 430 | | |
| 45 | | | 435 | | | | | 440 | | Gln | | | 445 | | | |
| | | 450 | | | | | 455 | | | Gly | | 460 | | | | |
| 50 | 465 | | | | | 470 | | | | Ala | 475 | | | | | 480 |
| | | | | | 485 | | | | | Asn 490 | | | | | 495 | |
| 55 | | | | 500 | | | | | 505 | Lys | | | | 510 | | |
| 33 | | | 212 | | | | | 520 | | Gln | | | 525 | | | |
| | | 530 | | | | | 535 | | | Ser | | 540 | | | | |
| 60 | 545 | | | | | 550 | | | | Thr | 555 | | | | | 560 |
| | | | | | 565 | | | | | Gly 570 | | | | | 575 | |
| 65 | | | | 580 | | | | | 585 | Ser | | | | 590 | | |
| 00 | | | 595 | | | | | 600 | | Ala | | | 605 | | | |
| | | 610 | | | | | 615 | | | Thr | | 620 | | | | |
| 70 | 625 | | | | | 630 | | | | Leu | 635 | | | | | 640 |
| | | | | | 645 | | | | | Lys 650 | | | | | 655 | |
| 7 5 | | | | 660 | | | | | 665 | Asn | | | | 670 | | |
| , , | тте | мта | AEP | rys | Phe | Met | Ala | Gly | Thr | Туr | Pro | Glu | Lys | Ala | Ala | Ilę |



```
Ala Ala Val Tyr Val Met Pro Ser Ala Pro Asp Ser Thr Phe His Leu
690 695 700
      Phe Leu Lys Ser Asn Thr Asn Arg Arg Leu Gln Lys Val Thr Thr Pro 705 710 715 720
      Ser Asp Trp Gln Ala Gly Thr Trp Leu Arg Ile Asn Leu Asp Lys Pro
725 730
      Phe Pro Val Asn Asn Asp His Met Leu Phe Ala Gly Ile Arg Met Pro 740 745 750
10
      Asn Lys Tyr Lys Leu Asn Arg Ala Ile Arg Tyr Val Arg Asn Pro Asp 755 760 765
      15
      Lys Tyr Leu Val Val Asn Thr Asp Ala Pro Lys Ile Asp Met Ser Leu 805 810 815
      Val Gln Glu Pro Tyr Ala Lys Gly Thr Asn Val Ala Pro Phe Pro Glu
820 825 830
      Leu Val Gly Ile Tyr Val Tyr Lys Asn Gly Thr Phe Ile Gly Thr Gln
835 840 845
20
      Asp Pro Ser Val Thr Thr Tyr Ser Val Ser Asp Gly Thr Glu Ser Asp
850 855 860
      Glu Tyr Glu Ile Lys Leu Val Tyr Lys Gly Ser Gly Ile Ser Asn Gly
865 870 875 880
25
      Val Ala Gln Ile Glu Asn Asn Asn Ala Val Val Ala Tyr Pro Ser Val
885 890 895
       Val Thr Asp Arg Phe Ser Ile Lys Asn Ala His Het Val His Ala Ala
900 905 910
      Ala Leu Tyr Ser Leu Asp Gly Lys Gln Val Arg Ser Trp Asn Asn Leu
915 920 925
30
       Arg Asn Gly Val Thr Phe Ser Val Gln Gly Leu Thr Ala Gly Thr Tyr
930 935 940
       Met Leu Val Met Gln Thr Ala Asn Gly Pro Val Ser Gln Lys Ile Val
35
       945
      Lys Gln
       (2) INFORMATION FOR SEQ ID NO:337
40
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 702 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
45
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
50
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
55
                  (B) LOCATION 1...702
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:337
       Trp Lys Lys Ser Asn Pro Gln Tyr Arg Gly Arg Thr Ala Asn Tyr Arg
1 5 10 15
60
      Asp His Tyr Asn Thr Glu Ser Tyr Leu Ser Leu Arg Arg Asn Gln Thr 20 25 30
      Net Lys Tyr Leu Ile Arg Leu Phe Leu Ser Leu Met Leu Leu Ser Leu 35 40
      Trp Thr Gly Cys Thr His Glu Glu Leu Ser Ile Cys Asp Gly Glu Asn 50 55
65
       Thr Leu Val Leu Arg Val Glu Thr Gly Lys Ala Pro Asn Ala Arg Ala 65 70 75 80
       Thr Glu Pro Gly Gln Gly Ile Tyr Asn Glu Asn Lys Val Gly Ser Ile 95 90 95
70
       Ser Val Leu Phe Tyr Leu Glu Gly Gln Leu Arg Trp Gln Val Lys Ser
100 105 110
```

Thr Asp Tyr Gln Ile His Glu Gly Ala Tyr Ile Ile Pro Val Lys Glu 115 120 125

Gln Het Arg Pro Leu Phe Asn Gly Asn Asn Asn Phe Ser Ile Tyr Val

75



| | | 130 Ala | Asn | Leu | Asp | | 135 Asn | Ala | Pro | Ala | | 140 Glu | Ala | Ala | Leu | |
|-----|------------|------------|------------|------------|------------|------------|------------|------------|-------------------|--------------------|------------|------------|------------|------------|------------|------------|
| 5 | 145 Gln | Phe | Val | Val | Glu 165 | 150 Lys | Ser | Ile | Glu | Val 170 | 155 Ser | Ser | Thr | Thr | Ala 175 | 160 Pro |
| J | Ala | Asp | Phe | Val 180 | | Leu | Ala | His | Gly 185 | | Lys | Gln | Ile | Asn 190 | | Ala |
| | Thr | Thr | Glu 195 | Gly | Lys | Leu | Leu | Gly 200 | Asp | Туг | Lys | Leu | Lys 205 | Arg | Val | Ala |
| 10 | | 210 | | | | | 215 | | Thr | | | 220 | | | | |
| | 225 | | | | | 230 | | | Phe | | 235 | | | | | 240 |
| 15 | | | | | 245 | | | | Ile | 250 | | | | | 255 | - |
| | | | | 260 | | | | | Glu 265 Trp | | | | | 270 | | |
| 20 | | | 275 | | | | | 280 | Lys | | | | 285 | | | |
| | - | 290 | | | | | 295 | | Туг | | | 300 | | | | |
| | 305 | | | | | 310 | | | Leu | | 315 | | | | | 320 |
| 25 | Glu | Ile | Leu | | 325 Ser | Leu | Gln | G1 u | Pro | 330 Glu | Ala | Val | Ser | | 335 Asn | Gly |
| | Thr | Leu | | 340 11e | Glu | Glu | Trp | | 345 Leu | His | Gln | Asp | | 350 Phe | Asn | Leu |
| 30 | Pro | Ala 370 | 355 Thr | Asn | Tyr | Leu | Ile 375 | 360 Val | Glu | Gln | His | Glu 380 | 365 Ile | Phe | Met | Asn |
| | Asn 385 | | Asn | Thr | Tyr | Ser 390 | | Lys | Tyr | Gln | Thr 395 | | Gln | Lys | Pro | Ile 400 |
| 35 | | Ile | Ser | Ile | Gln 405 | | Val | Thr | Phe | Ser 410 | | Val | Ser | Ser | Asp 415 | |
| | Thr | Gln | His | Asn 420 | Asp | Leu | Val | Ala | Ser 425 | Ser | Ser | Asp | Gln | Tyr 430 | Pro | Thr |
| 40 | | | 435 | | | | | 440 | Ile | | | | 445 | | | |
| 40 | | 450 | | | | | 455 | | Phe | | | 460 | | | | |
| | 465 | | | | | 470 | | | Leu | | 475 | | | | | 480 |
| 45 | | | | | 485 | | | | Ile | 490 | _ | | | | 495 | |
| | | | | 500 | | | | | 505 Gly | | | | | 510 | | |
| 50 | | | 515 | | | | | 520 | His | | | | 525 | | | |
| | | 530 Ile | Vāl | Trp | Ser | | 535 Thr | Asn | Thr | Thr | | 540 Gln | Asp | Leu | Glu | |
| 55 | 545 Ser | Arg | Met | Ile | | 550 Pro | Ser | Phe | Glu | | 555 Ala | Ser | Gln | Leu | | 560 Ala |
| 33 | Thr | Leu | Pro | Met 580 | 565 Pro | Tyr | Leu | Glu | Tyr 585 | 570 Trp | Pro | Gly | Thr | Ser 590 | 575 Tyr | Leu |
| | Leu | Asp | Tyr 595 | | G1 y | Asn | Tyr | Asn 600 | Asn | Lys | Arg | туг | Ala 605 | | Phe | Asn |
| 60 | C?.2 | Ala 610 | Phe | Tyr | Trp | Glu | Lys 615 | | Lys | Val | Asn | Asn 620 | | Glu | Ile | Lys |
| | Phe 625 | Asp | Asp | Trp | Arg | Leu 630 | Pro | Thr | Glu | Ala | Glu 635 | Ile | Lys | Leu | Ile | Asp 640 |
| 65 | - | | | | 645 | | | | Ala | 650 | | | | | 655 | - |
| | | | | 660 | | | | | Ala 665 | | | | | 670 | | |
| 70 | _ | _ | 675 | _ | | _ | | 680 | Ser | _ | | - | 685 | - | cys | val |
| , 0 | vrd | 690 | vaı | ьуз | пλε | LTO | 695 | wig | Asp | υλ _, ε, | гуg | 700 | GIA | ኮንደ | | |
| | (2) | INF | ORMA | rion | FOR | SEQ | ID 1 | 10:3 | 38 | | | | | | | |

75 (i) SEQUENCE CHARACTERISTICS:



```
(A) LENGTH: 1312 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 5
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
10
                   (A) ORGANISM: Perphyromenas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...1312
15
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:338
       20
       Ile Cys Val Ile Gln Lys Phe Thr Asn Val Lys Leu Asn Asp Met Arg 20 \hspace{1.5cm} 25 \hspace{1.5cm} 30 \hspace{1.5cm}
       Lys Ile Leu Ser Phe Leu Met Met Cys Ser Leu His Leu Gly Leu Gln 35 40 45
       Ser Gln Thr Trp His Gly Asp Pro Asp Ser Val Ala Ala Leu Pro Ser 50 55 60
25
       Ile Gly Ile Gln Glu Ser Ser Cys Thr Arg Ile Thr Phe Glu Val Val
65 70 75 80
       Phe Pro Gly Phe Tyr Ser Val Glu Lys Arg Glu Gly Asn Gln Val Phe
85 90
30
       Gln Arg Ile Ser Het Pro Gly Cys Gly Ser Phe Gly Asn Leu Gly Glu
100 105 110
       Ala Glu Leu Pro Val Leu Lys Lys Met Ile Ala Val Pro Glu Phe Ser
115 120 125
       Thr Ala Asn Val Ala Val Lys Ile Lys Glu Thr Glu Thr Phe Asp Asn
130 135 140
35
       Tyr Asn Ile Tyr Pro Asn Pro Thr Tyr Val Val Glu Glu Leu Pro Glu
145 150 155 160
       Gly Gly Thr Tyr Leu Val Glu Ala Phe Ala Ile Asn Asn Asp Tyr Tyr
165 170 175
       Ser Gln Asn Val Ser Leu Pro Ser Thr His Tyr Val Tyr Ser Gln Asp
180 185 190
40
       Gly Tyr Phe Arg Ser Gln Arg Phe Ile Glu Val Thr Leu Tyr Pro Phe
195 200 205
       Arg Tyr Asn Pro Val Arg Gln Glu Ile Leu Phe Ala Lys Lys Ile Glu 210 215 220
45
       Val Thr Ile Thr Phe Asp Asn Pro Gln Pro Pro Leu Gln Lys Asn Thr
225 230 235 240
       Gly Ile Phe Asn Lys Val Ala Ser Ser Ala Phe Ile Asn Tyr Glu Ala
245 250 255
50
       Asp Gly Lys Ser Ala Ile Glu Asn Asp Met Val Phe Ser Arg Gly Thr 260 265 270
       Thr Thr Tyr Ile Ser Gly Asn Val Ala Ser Asn Leu Pro Gln Asn Cys 275 280 285
       Asp Tyr Leu Val Ile Tyr Asp Asp Met Phe Asn Val Asn Gln Gln Pro 290 295 300
55
       His Asp Glu Ile Lys Arg Leu Cys Glu His Arg Ala Phe Tyr Asn Gly 305 310 315 320
       Phe Asp Val Ala Ala Val Ser Ile Lys Asp Val Leu Asn Ser Phe Pro 325 330 335
60
       Ser Asn Ala Thr Ser Tyr Ile Asn Glu Thr Lys Leu Lys Asn Phe Ile 340 345 350
       Arg Ser Val Tyr Asn Gln Ser Asn Ala Lys Arg Thr Leu Asp Gly Lys
355 360 365
       Leu Gly Tyr Val Leu Leu Ile Gly Lys Pro Leu Ser Lys Tyr Leu Ala 370 375 380
65
       Asp Thr Asp Asn Thr Lys Val Pro Thr Ser Phe Ile His Asn Val Ser 385 390 395 400
       Leu Ile Pro Ser His Pro Thr Phe Gly Ser Ile Cys Ala Ser Asp Tyr
405 410 415
70
       Phe Phe Ser Cys Val Ser Pro Leu Asp Thr Val Gly Asp Leu Phe Ile
420 425 430
       Gly Arg Phe Ser Val Thr Asn Ala His Glu Leu His Asn Leu Ile Glu
435 440 445
       Lys Thr Ile Asn Lys Glu Ile Ser Tyr Asn Pro Ile Ala His Lys Asn
75
```

| | Ile 465 | Leu | Tyr | Ala | Glu | Gly 470 | | Gly | Cys | Asp | | Pro | Ile | Leu | Arg | Leu 480 |
|------------|------------|------------|------------|------------|-------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|-------------|--------------------|
| | | Leu | Lys | Glu | | | | Gly | Tyr | Thr | 475 Val | Asn | Ser | Ile | | |
| 5 | Ser | Asn | Gln | Val | 485 Ser | Ala | Ile | Asp | Ser | 490 Ile | Phe | Asp | Cys | Leu | 495 Asn | Asn |
| | Gly | Ser | His | 500 His | Phe | Tyr | Phe | Asn | 505 Thr | His | Gly | Met | Pro | 510 Thr | Vāl | Trp |
| | Gly | Ile | 515 Gly | Gln | Gly | Leu | Asp | 520 Val | Asn | Thr | Leu | Thr | 525 Ala | Arg | Leu | Asn |
| 10 | Asn | 530 Thr | Ser | Ser | Gln | | 535 Leu | Cys | Thr | Ser | Leu | 540 Ser | Cys | Ser | Ser | Ala |
| | 545 Val | Ala | Asp | Ser | Thr | 550 Ile | Arg | Ser | Leu | Gly | 555 Glu | Val | Leu | Thr | Thr | 560 Tyr |
| 15 | Ala | Pro | Asn | | 565 Glγ | Phe | Ser | Ala | Phe | 570 Leu | Gly | Gly | Ser | Arg | 575 Ala | Thr |
| | Gln | Tyr | Ala | 580 Val | Tyr | Leu | Glu | Gly | 585 Pro | Суѕ | Pro | Pro | Ser | 590 Glu | Phe | Туг |
| | Glu | Tyr | 595 Leu | Pro | Tyr | Ser | Leu | 600 Tyr | His | Asn | Leu | Ser | 605 Thr | Val | Val | Gly |
| 20 | | 610 | | | | | 615 | | | Thr | | 620 | | | | |
| | 625 | | | | | 630 | | | | | 635 | | | - | | 640 |
| 25 | | | | | 645 | | | | | Asp 650 | | | | | 655 | |
| 25 | | | | 660 | | | | | 665 | Ile | | | | 670 | | |
| | | | 675 | | | | | 680 | | Asn | | | 685 | | | |
| 30 | Pro | Glu 690 | Γ?.ε | Gly | Val | Leu | His 695 | Phe | Thr | Asn | Asn | Gly 700 | Ser | Ile | Gln | Va1 |
| | Het 705 | Ser | Gly | Gly | Thr | Leu 710 | Glu | Ile | Gly | Asn | Gln 715 | Ala | Lys | Ile | Ser | Gly 720 |
| | Glu | Thr | Glγ | Ala | Asn 725 | Pro | Thr | Phe | Ile | Thr 730 | | Tyr | Gly | Asp | Gly 735 | Leu |
| 35 | Ala | lle | Asn | Lys 740 | Gln | Val | Glu | Ile | Asp 745 | Asn | Ile | Asp | Arg | Leu 750 | Asn | Leu |
| | Phe | Ser | Thr 755 | His | Ser | Val | Met | Pro 760 | Lys | Phe | His | Phe | Asp 765 | Ser | Val | Lys |
| 40 | Phe | Asn 770 | Ser | Ala | Pro | Leu | Туг 775 | Thr | Thr | Asn | Сув | 11e 790 | Val | Glu | Ile | Ser |
| | 785 | | | | | 790 | | | | Ile | 795 | | | | | 800 |
| | Leu | Ser | Val | Glu | Asn 805 | Ser | Met | Phe | Ser | Ser 810 | Ser | Gly | Ile | Thr | Val 815 | Phe |
| 4 5 | Lys | Pro | Met | Ala 820 | Thr | Ser | Ser | Ile | Thr 825 | Glу | Leu | Ser | Thr | Lys 830 | Ala | Lys |
| | Ile | Thr | Asp 835 | Asn | Thr | Phe | Phe | Ala 840 | Thr | G1 Y | Asn | Phe | Ala 845 | Tyr | Hi <i>s</i> | Ile |
| 50 | | 850 | | | | | 855 | | | Ser | | 860 | | | _ | |
| | Asp 865 | Asn | IJ€ | Pro | Glu | Tyr 870 | Tyr | Ile | Ser | Gl. | A≘n 875 | Lys | I).e | Val | Asn | Су <i>є</i> 880 |
| | Asp | Glu | Alā | Leu | Val 885 | Leu | Asn | Asn | Ser | G1y 890 | Asn | Arg | Thr | Asn | Arg 895 | Leu |
| 55 | His | Asn | Ile | Thr 900 | Arg | Asn | Val | | Lys 905 | Asn | Cys | Arg | Ile | Gly 910 | | Thr |
| | | | 915 | | | | | 920 | | Arg | | | 925 | Ser | | |
| 60 | His | 11e 930 | G17. | Vāl | Arg | Leu | Leu 935 | Asn | Asn | Ser | c),s | Phe 940 | туг | Phe | Asp | Asn |
| | Ala 945 | Pro | Val | Ile | Asn | Glu 950 | Glu | Asp | Lys | Gln | Thr 955 | Phe | Ile | Ser | Asn | Arg 960 |
| | Thr | Trp | Glrı | Leu | Tyr 965 | Ser | Ser | Asn | Gly | Thr 970 | Phe | Pro | Leu | Asn | Phe 975 | His |
| 65 | Tyr | Asn | Ser | Leu 980 | Gln | Gly | Gly | Asp | Thr 985 | Asp | Thr | Trp | Ile | Tyr 990 | Asn | Aεp |
| | Thr | Tyr | Thr 995 | Asn | Arg | Туг | Ile | Asp 1000 | Val | Ser | Asn | Asn | His 100 | Trp | Gly | Asn |
| 70 | | 1010 |) | | | | 101 | 5 | | Phe | | 10 | Pro 20 | Asp | | |
| | 1025 | 5 | | | | 103 | Asp 80 | Gly | | Pro | 10 | Gly 35 | Arg | | | 1040 |
| | Ser | Ser | Ala | Glu | Ala 1045 | | Glu | Phe | Gln | Thr 105 | Ala | | Asp | Суѕ | | |
| 75 | Asn | Ser | Asp | Tyr | | | λla | Lys | Val | Ala | | Lys | Met | Met | | |

| | | | | 1060 | ^ | | | | 10 | 65 | | | | , | 070 | |
|------------|---|-------------|---------------|-------|----------------|----------|-----------|------------|-------|-------|------|------|-----|------|------------|------|
| | Thr | Tyr | Pro 107 | Glu | Ser | Asp | Phe | Ala 108 | | | Ala | Leu | | | 070 Leu | Phe |
| 5 | Arg | Ile 1090 | Glu | | Met | Ser | Gly 10 | Asn | | Туг | Glu | | | | Asp | Tyr |
| | 110 | 5 | | | Pro | 111 | 10 | | | | 11 | 115 | | | | 1120 |
| 40 | | | | | Ser 1125 | 5 | | | | 113 | 30 | | | | 1. | 135 |
| 10 | | | | 1140 | | | | | 114 | 15 | | | | 1 | 150 | - |
| | | | 1159 | 5 | Phe | | | 116 | 50 | | | | 1: | 165 | | |
| 15 | | 1170 |) | | Ser | | 11 | 75 | | | | 1: | 180 | | | |
| | 118 | 5 | | | Arg | 119 | 90 | | | | 1.1 | 95 | | | _ | 1200 |
| 20 | | | | | Thr 1205 | . | | | | 121 | LO | | | | 1: | 215 |
| 20 | | | | 1220 | | | | | 122 | 25 | | | | 1: | 230 | |
| | | | 1235 | 5 | Ala | | | 124 | 10 | | | | 12 | 245 | | • |
| 25 | | 1250 |) | | Ser Gly | | 125 | 55 | | | | 12 | 260 | | | |
| | 126 | 5 | | | Thr | 127 | 70 | | | | 12 | 275 | _ | _ | _ | 1280 |
| 30 | | | | | 1285 Gln | 6 | | | | 129 | 90 | | _ | | 12 | 295 |
| | | | | 1300 |) | - | | | 130 | | | | | | 310 | -1- |
| 05 | (2) | INFO | ORMAT | MOIT | FOR | SEQ | ID ! | 10:33 | 39 | | | | | | | |
| 35 | 35 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 938 amino acids (B) TYPE: amino acid | | | | | | | | | | | | | | | |
| 40 | | | | | OPOLO | | | | | | | | | | | |
| 40 | | | | | E TY | | - | ein | | | | | | | | |
| | | | | | ETICA | | | | | | | | | | | |
| 45 | | (V1) | | | AL SC RGANI | | | hyro | mona | ıs gi | ngiv | alis | 5 | | | |
| | | (ix) |) FE <i>F</i> | | : ME/F | EY: | misc | : fea | ature | 9 | | | | | | |
| 50 | | | (E | 3) LC | CATI | .01 | 9 | 38 | | | | | | | | |
| | | | | | CE DE | | | | | | | | | | | |
| = = | 1 | | | | Arg 5 | | | | | 10 | | | | | 15 | |
| 55 | | | _ | 20 | Ile | | | _ | 25 | | | | | 30 | | |
| | | | 35 | | Thr | | | 40 | | | | | 45 | | | |
| 60 | | 50 | | | Ala | | 55 | | | | | 60 | | | | |
| | 65 | | | | Gly | 70 | | | | | 75 | | | | | 80 |
| 65 | | | | | Ala 85 | | | | | 90 | | | | | 95 | |
| 00 | | | | 100 | Asn | | | | 105 | | | | | 110 | | |
| | | | 115 | | Gly Tyr | | | 120 | | | | | 125 | | | |
| 70 | | 130 | | | Met | | 135 | | | | | 140 | | | | |
| | 145 | | | | Asp | 150 | | | | | 155 | | | | | 160 |
| <i>7</i> 5 | | | | | 165 | | | | | 170 | | | | | 175 | |
| | ττb | my | nry | nr 9 | lle | val | nεp | n-cu | r r o | GIU | ar? | 1111 | гÀЕ | 12.1 | TTE | wra |



| | | | | 180 | | | | | 185 | | | | | 190 | | |
|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Trp | Arg | His 195 | Туr | Lys | Val | Thr | Asp 200 | | His | Thr | Glu | Phe 205 | | Lys | Leu |
| 5 | Asp | Asp 210 | | Thr | Val | Tyr | Arg 215 | Ser | | Glu | Gly | Pro 220 | | | Ala | Thr |
| | Asp 225 | Phe | Thr | Val | Ile | Asn 230 | Ile | Gly | Gln | Asn | Val 235 | Gly | Arg | Leu | Thr | Trp 240 |
| | Asn | Tyr | Pro | Glu | Asp 245 | | Gln | Pro | Glu | Gly 250 | Γλε | Gly | Asn | Glu | Glu 255 | |
| 10 | Gln | Leu | Ser | Gly 260 | Туr | Asn | Ile | Tyr | Ala 265 | Asn | Gly | Thr | Leu | Leu 270 | | Gln |
| | Ile | Lys | Asp 275 | Val | Ser | Ile | Leu | Glu 280 | Tyr | | Asp | Ser | Thr 285 | | Ser | Leu |
| 15 | Arg | Asp 290 | Asn | Pro | Leu | Gln | Val 295 | Glu | Туг | Суѕ | Val | Thr 300 | | Val | Tyr | Asp |
| | Glu 305 | Ser | lle | Glu | Ser | Ser 310 | Thr | Val | Суѕ | Glу | Thr 315 | Leu | His | Tyr | Ala | Thr 320 |
| | Asp | Ala | lle | Leu | Tyr 325 | Glu | Asn | Phe | Glu | Asn 330 | Gly | Pro | Val | Pro | Asn 335 | Gly |
| 20 | Trp | Leu | Va1 | 11e 340 | | Ala | Asp | Glγ | Asp 345 | | Phe | Ser | Trp | Gly 350 | | Tyr |
| | Leu | Asn | Ala 355 | | Asp | Alā | Phe | Pro 360 | | His | Asn | Gly | Gly 365 | | Сув | Ser |
| 25 | Leu | Ser 370 | | Ser | Туr | Val | Pro 375 | | Ile | Gly | Pro | Val 380 | | Pro | Asp | Asn |
| | Tyr 385 | Leu | lle | Thr | Pro | Lys 390 | Vāl | Glu | Gly | Ala | Lys 395 | | Val | Lys | Tyr | Trp |
| | Val | Ser | Thr | Gln | Asp 405 | Ala | Asn | Trp | Ala | Ala 410 | Glu | His | Tyr | Ala | Val 415 | Het |
| 30 | Ala | Ser | Thr | Thr 420 | Gly | Thr | Ala | Val | Gly 425 | Asp | Phe | Väl | Ile | Leu 430 | Phe | Glu |
| | | | 435 | | | | | 440 | | | Trp | | 445 | | | |
| 35 | | 450 | | | | | 455 | | | | Trp | 460 | | | | |
| | Thr 465 | Asp | Ile | Tyr | Phe | Leu 470 | r7.2 | Leu | Asp | Asp | Ile 475 | Thr | Val | Phe | Gl y | Thr 480 |
| | Pro | Ala | Ser | Glu | Pro 485 | Glu | Pro | Val | Thr | Asp 490 | Phe | Val | Val | Ser | Leu 495 | Ile |
| 40 | Glu | Asn | Asn | Lys 500 | Gly | Arg | Leu | Lys | Trp 505 | Asn | Tyr | Pro | Asn | Gly 510 | Tyr | Glu |
| | | | 515 | | | | | 520 | | | Gln | | 525 | _ | = | |
| 45 | | 530 | | | | | 535 | | | | Ile | 540 | | | | |
| | 545 | | | | | 550 | | | | | Arg 555 | | | | | 560 |
| 50 | | | | | 565 | | | | | 570 | Asp | | | | 575 | |
| 50 | | | | 580 | | | | | 585 | | Gln | | | 590 | | |
| | | | 595 | | | | | 600 | | | Glu | | 605 | | | |
| 55 | | 610 | | | | | 615 | | | | | 620 | | | | |
| | 625 | | | | | 630 | | | | | Ser 635 | | | | | 640 |
| CO | | | | | 645 | | | | | 650 | Leu | | | | 655 | |
| 60 | | | | 660 | | | | | 665 | | Ser | | | 670 | | |
| | | | 675 | | | | | 680 | | | Ser | | 685 | | | |
| 65 | | 690 | | | | | 695 | | | | Thr | 700 | | | | |
| | 705 | | | | | 710 | | | | | Leu 715 | | | | | 720 |
| 70 | | | | | 725 | | | | | 730 | Asp | | | | 735 | |
| 70 | | | | 740 | | | | | 745 | | Glu | | | 750 | | |
| | | | 755 | | | | | 760 | | | Asn | | 765 | | | |
| 75 | Lys | Trp 770 | Asn | Tyr | Pro | Asn | Gly 775 | Tyr | Glu | Pro | Asp | Lys 780 | Thr | Asp | Asp | Lys |
| | | | | | | | | | | | | | | | | |



```
Lys Pro Leu Gln Leu Thr Gly Tyr Asn Ile Tyr Ala Asn Gly Ser Leu
785 790 795 800
       Leu Val His Ile Gln Asp Pro Thr Val Leu Glu Tyr Ile Asp Glu Thr
805 810 815
                        805
                                               810
       Tyr Ser Ser Arg Asp Gly Gln Val Glu Met Glu Tyr Cys Val Thr Ala
820 825 830
       Val Tyr Asn Asp Asn Ile Glu Ser Gin Ser Val Cys Asp Lys Leu Asn
835 840 845
       Tyr Thr Ile Thr Ser Leu Asp Asn Ile Gln Ser Asp Thr Ser Leu Lys
850 855 860
10
       Ile Tyr Pro Asn Pro Ala Ser Tyr Val Val Arg Ile Glu Gly Leu Ser
865 870 875 880
       Arg Ser Lys Ser Thr Ile Glu Leu Tyr Asn Ala Leu Gly Ile Cys Ile
885 890 895
       Leu Arg Glu Glu Thr His Ser Glu Lys Thr Glu Ile Asp Val Ser Arg
900 905 910
15
       Leu Asn Asp Gly Val Tyr Leu Ile Lys Val Val Gly Gly Asn Lys Thr
915 920 925
       Thr Thr Glu Lys Val Glu Ile Lys Arg Pro
20
       (2) INFORNATION FOR SEQ ID NO:340
             (i) SEQUENCE CHARACTERISTICS:
25
                   (A) LENGTH: 606 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
30
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
35
           (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...606
40
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:340
       Ile Ile Phe Cys Thr Ile His His Ser Glu Leu Glu Ile Met Asn Ser 1 5 10 15
       Ile Met Lys Tyr Gln Leu Tyr Thr Ala Val Ile Met Ala Leu Ser Val 20 25 30
45
       Ser Ser Val Cys Gly Gln Thr Pro Arg Asn Thr Glu Thr Lys Arg Pro 35 40 45
       Asp Thr Leu Arg Arg Glu Leu Thr Ile Val Asn Asp Gln Thr Val Glu
50 60
       Net Glu His Ala Asp Pro Leu Pro Ala Ala Tyr Lys Ala Ile Glu Pro 65 70 75 80
50
       Arg Leu Lys Pro Phe Arg Pro Glu Tyr Asn Lys Arg Thr Phe Gly Phe 85 90 95
       Val Pro Glu Val Ser Ser Ser Gly Arg Asn Asn Leu Pro Asn Ile Leu
100 105 110
55
       Pro Thr Glu Gly His Net Lys His Arg Gly Tyr Leu Asn Ile Gly Ile
115 120 125
       Gly His Thr Leu Asn Gln Arg Met Asp Ala Gly Tyr Arg Leu Ile Asp
130 135 140
       Ala Glu Gln Glu Arg Leu Asn Leu Phe Leu Ser Tyr Arg Gly Met Lys
145 150 155 160
60
       Ser Ala Phe Asn Thr Gly Asp Phe Asp Gly Asp Arg Lys Asp Arg Arg
165 170 175
       Met Net Ala Gly Val Asp Tyr Glu Gln Arg Arg Pro Ser Phe Val Leu
180 185 190
65
       Ala Thr Gly Leu Tyr Tyr Ser Asn His Tyr Phe Asn Asn Tyr Gly Arg
195 200 205
       Gly Ala Thr Thr Asn Val Gly Ser IIe Pro Gln Leu Ser Thr Pro Val
210 215 220
70
       Thr Pro Gln Met Asp Asn Gly Thr His Asn Val Arg Val Tyr Leu Gly 225 230 235 240
       Ala Lys Asn Asp Val Ile Asp Ala Arg Ile Asp Tyr Arg Phe Phe Arg
245 250 255
       Ser Ile Pro Tyr Leu Gly Thr Asp Pro Met Lys Ala Leu Thr Glu His
75
```

```
Thr Pro Glu Leu Asn Val Thr Met Ser Asn Glu Leu Ser Asp Asp Ile
                275
                                        280
                                                               285
       Lys Leu Gly Val Glu Val Arg Thr Gly Gly Leu Phe Phe Ala Lys Asn
          290
                                 295
                                                           300
       Ser Glu Met Ile Gln Thr Gly Val Leu Ser Glu Thr Asp Arg Asn Leu 305 310 315 320
       Tyr Tyr Val Glu Gly Ala Pro Thr Ile Gly Phe Val Gly Asp Ser Asp 325 339 335
       As n Met Gln Trp As n Ile Gln Ala Gly Val Gly Ile Ser Ser His Phe 340 345 350
10
       Gly Ala Lys Gly Arg Leu Phe Phe Trp Pro Lys Leu Asp Ala Ser Leu 355 360 365
       Ser Ile Phe Pro Ser Trp Arg Val Tyr Ala Lys Ala Phe Gly Gly Val 370 375 380
       Ile Arg Asn Gly Leu Ala Asp Val Met Gln Glu Glu Met Pro Tyr Leu
385 390 395 400
15
       Het Pro Asn Thr Ile Val Leu Pro Ser Arg Asn Ala Leu Thr Ala Gln
405 415
       Leu Gly Val Lys Gly Asn Ile Ala Asp Val Val Arg Met Glu Val Tyr
420 430
20
       Gly Asp Phe Ser Lys Leu Thr Gly Val Pro Phe Tyr Thr Pro Thr Leu
435 440 445
       Pro Leu Tyr Asn Pro Ser Asp Leu Tyr Gln Tyr Asn Val Ser Phe Leu
450 455 460
       Pro Ile Tyr Ala Asp Gly Ser Arg Trp Arg Ala Gly Gly Lys Leu Glu
465 470 470 480
Tyr Ser Tyr Arg Asp Met Leu Arg Phe Leu Val Asp Ala Ser Tyr Gly
485 485 480 490 495
25
       Lys Trp Asn Leu Asp Gly Gly Leu Val Ala Ser Met Gln Pro Asp Leu 500 505 510
30
       Ile Leu Lys Ala Glu Val Gly Val His Pro Ile Ala Pro Leu Asp Val 515 520 525
       Arg Leu Arg Tyr Thr Gln Leu Asn Gly Arg Tyr Arg Tyr Ser Phe Gly 530 535 540
35
       Ser Ala Gly Ser Glu Ala Leu Gly Ile Gly Asn Val His Leu Leu Ser
545 550 560
       Ala Asp Val Ser Tyr Lys Leu Lys Lys Asn Leu Ser Leu Tyr Leu Lys 565 570 575
       Ile Asp Asn Net Leu Ala Glu Thr Thr Glu Leu Ile Gly Tyr Tyr Pro
40
       Met Gln Pro Phe His Cys Phe Ala Gly Phe Ser Trp Thr Phe
                                       600
       (2) INFORMATION FOR SEQ ID NO:341
45
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 357 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
50
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
55
            (vi) ORIGINAL SOURCE:
                   (A) ORGAHISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
60
                   (B) LOCATION 1...357
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:341
       Asn Ser His Ala Thr Val Ser Val Ile Cys Ser Met Het Glu Lys Cys
1 5 10
65
       Ile Phe Ala His Tyr Pro His Asn Leu Val Phe Met Ile Arg Lys His
                                            25
                                                                  30
       Phe Gly Ile Ile Leu Gly Phe Leu Ser Leu Val Phe Ser Ala Gly Ala
35 40 45
70
       Gln Gln Glu Lys Gln Val Phe His Phe Leu Asn Leu Pro Ala Thr Ala
50 55 60
       Gln Ala Leu Ala Ala Gly Gly Lys Ala Ile Thr Ile Val Asp Asp Asn
65 70 75 80
       Pro Gly Leu Ala Phe Glu Asn Pro Ala Leu Leu Gly Tyr Glu Ser Gly
75
```

```
Gly Asn Ala Cys Tyr Ala Ser Ser Val Gly Glu Arg Gly Met Trp Gly
115 120 125
        Val Gly Met Arg Phe Leu Asn Tyr Gly Ser Met Gln Gly Tyr Asp Gln 130 135 140
        Asn Ala Ile Ala Thr Gly Ser Phe Ser Ala Ser Asp Ile Ala Val Gln
145 150 150 160
        Gly Phe Tyr Ser His Glu Leu Ser Asn His Phe Arg Gly Gly Val Ser
165 170 175
10
        Leu Lys Ala Leu Tyr Ser Ser Ile Glu Thr Tyr Ser Ser Phe Gly Leu 180 185 190

Gly Val Asp Val Gly Ile Ser Tyr Tyr Asp Asp Asp Lys Gly Tyr Ser 195 200 205
        Ala Ser Ala Leu Phe Lys Asn Val Gly Ala Gln Leu Lys Gly Tyr Asn
210 215 220
15
        Glu Glu Arg Glu Pro Leu Asp Trp Asp Phe Gln Leu Gly Phe Ser Arg
225 230 235 240
                                                          235
        Ser Phe Ile Asn Ala Pro Phe Arg Leu His Ile Thr Leu Phe Asn Leu
245 250 255
20
        Asn Pro His Tyr Phe Lys Arg Leu Val Pro Arg Asp Leu Ser Lys Met
260 265 270
        Gln Lys Phe Leu Arg His Phe Ser Ile Gly Ala Glu Phe Thr Pro Ser
275 280 285
        Glu Arg Phe Trp Val Gly Leu Gly Tyr Thr Pro Gln Ile Ala Gln Asp
290 295 300
25
        290 295 300

Phe Glu Val Glu Gly Gly Asn Lys Trp Gly Gly Leu Ser Ala Gly Val 305 315 320

Gly Phe Thr Ser Gly Val Val Arg Val Gly Val Ser Ala Ala Thr Tyr 325 330 335

His Pro Ala Ala Leu Ser Phe Met Cys Ser Val Gly Ile Arg Leu Asp 340 345
30
        Asp Lys Ser Ile Phe
35
         (2) INFORMATION FOR SEQ ID NO: 342
               (i) SEQUENCE CHARACTERISTICS:
                      (A) LENGTH: 337 amino acids
40
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
              (ii) MOLECULE TYPE: protein
45
             (iii) HYPOTHETICAL: YES
              (vi) ORIGINAL SOURCE:
                      (A) ORGANISM: Perphyromenas gingivalis
50
              (ix) FEATURE:
                      (A) NAME/KEY: misc_feature
                      (B) LOCATION 1...337
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:342
55
         Pro Gln Pro Val Gly Leu Lys Glu Ile Thr Tle Lys Pro Het Cys Leu
1 5 10 15
         Glu Pro Ile Ile Ala Pro Ile Ser Ser Glu Leu Leu Glu Gln Glu Leu
20 25 30
60
         Thr Ala Asp Arg Phe Leu Arg Met Thr Asn Lys Ala Gly Asn Glu Ile
35 40 45
         Tyr Val Phe Thr Ala Glu Glu Ala Pro His Cys Met Lys Glu Val Gly 50 60
         Arg Leu Arg Glu Glu Ala Phe Arg His Tyr Gly Gly Gly Thr Gly Lys
65 70 75 80
65
        Ala Ile Asp Ile Asp Glu Phe Asp Thr Net Pro Gly Ser Tyr Lys Gln
85 90 95
        Leu Ile Val Trp Asp Pro Gln Asn Lys Ala Ile Leu Gly Gly Tyr Arg
100 105 110

Phe Ile Tyr Gly Arg Asp Val Ala Phe Asp Thr Asp Gly Lys Pro Leu
115 120 125
70
         Leu Ala Thr Ala Glu Met Phe Arg Phe Ser Asp Ala Phe Leu His Asp
130 135 140
         Tyr Leu Pro Tyr Thr Val Glu Leu Gly Arg Ser Phe Val Ser Leu Gln
145 150 155 160
75
```

Gly Arg Ala Phe Leu Ser Tyr Leu Tyr Tyr Het Ser Gly Ser His Met $100 \hspace{1cm} 105 \hspace{1cm} 110$



| | Tur G | ln Ser | Thr | Ara | Met | Glv | Thr | Lvs | Ala | Tle | Phe | Val | Leu | Aen | Aen |
|----------------------------|---|--|--|--|--|--|--|--|---|--|--|---|--|---|---|
| | | | | 165 | | | | | 170 | | | | | 175 | |
| _ | Leu T | | 180 | | | | | 185 | | | | | 190 | | |
| 5 | Tyr P | ne Tyr 195 | | Lys | Val | Thr | Met 200 | Tyr | Lys | Asp | Tyr | Aεp 205 | Arg | Arg | Ala |
| | Arg As | | Ile | Leu | Tyr | Phe 215 | | Arg | Lys | His | Phe 220 | Ser | Asp | Pro | Glu |
| 10 | Gly Le 225 | | Lys | Pro | Ile 230 | | Pro | Leu | Pro | Ile 235 | | Ile | Ser | Ala | Glu 240 |
| | Asp G | u Ala | Leu | Phe 245 | | Ser | Ser | Asp | Phe 250 | | Thr | Asn | Tyr | Lys 255 | |
| | Leu As | n Ile | Glu 260 | | Arg | ГЛЗ | Leu | Gly 265 | | Asn | Ile | Pro | Pro 270 | | Val |
| 15 | Ser Al | a Tyr 275 | Ile | Ala | Leu | Ser | Pro 280 | | Met | Arg | Val | Phe 285 | | Thr | Ala |
| | Val As | n Glu | | Phe | Gly | Glu 295 | | Glu | Glu | Thr | Gly 300 | _ | Phe | Ile | Ala |
| 20 | Val G | | Ile | Leu | | | Lys | Lys | Gln | | | Ile | Glu | Ser | |
| 20 | 305 Ile Le | u Ser | Arg | | 310 Glu | L}:s | Lys | Gly | | 315 Asp | Ser | Ser | Asn | | 320 Arg |
| | Ser | | | 325 | | | | | 330 | | | | | 335 | |
| 25 | | | | | | | | | | | | | | | |
| | (2, I) | FORM | TIOH | FOR | SEQ | ID I | 10:3 | 13 | | | | | | | |
| | (| i) SE | | | | | | | | | | | | | |
| 30 | | { | | YPE: | amir | no ac | cid | acio | is | | | | | | |
| | | (| D) T(| OPOLO | OGY: | line | ear | | | | | | | | |
| | (i | i) MO | LECUI | LE T | YPE: | prot | ein | | | | | | | | |
| 35 | (ii | i) HY | POTH | ETICA | AL: Y | ES | | | | | | | | | |
| | (7 | i) OR | IGINA A) OI | | | | hure | omo n | | nais | | | | | |
| | | , | A) () | (GAIV. | LON. | FOT | ліуті | שונטוונ | as ui | | | | | | |
| 40 | , , | | N. M. J. D. F | ٠. | | | | | | 9. | | | | | |
| 40 | (i | | A) NA | AHE/I | | | | | | 9. | | | | | |
| 40 | (i | (| | AHE/I | | | | | | 9. | | | | | |
| 40 45 | | (| A) NA B) LO | AME/F CAT | ION 1 | 15 | 566 | ature | e | | | | | | |
| | | ((i) SE | A) NA B) LO | AME/H OCATI CE DE | EON 1 | PTIC | 566 DN: 5 | ature | e [D NO Leu |): 343 | 3 | | Arg | | Asp |
| | () | (i) SE | A) NA B) LO QUENO Ala Met | AME/F DCATI CE DE Arg 5 | [ON] ESCR] Leu | PTIC | 666 DN: 5 Tyr | SEQ I Thr Arg | ED NO Leu 10 |):343 Phe | 3 Arg | Asn | Pro | 15 | _ |
| | Tyr As | (i) SE p Gly s Pro | A) NA B) LO QUENO Ala Met 20 | AME/H DCATI CE DE Arg 5 Lys | ION 1 ESCRI Leu Thr | PTIC Val Ile | FG6 DH: S Tyr Val Ala | SEQ 1 Thr Arg 25 | ED NO Leu 10 Tyr | D:343 Phe Ser | B Arg Arg | Asn Leu Val | Pro 30 | 15 Val | Ala |
| 4 5 | Tyr As 1 1 | (i) SE p Gly s Pro le Phe 35 u Asn | A) NA B) LO QUENO Ala Met 20 Cys | AME/I DCATI CE DE Arg 5 Lys Leu | ION 1 ESCRI Leu Thr Leu | PTIC Val Ile Gly Glu | Fig. 566 Tyr Val Ala | SEQ 1 Thr Arg 25 Val | Leu 10 Tyr | D:343 Phe Ser Leu | Arg Arg Ser Ile | Asn Leu Val 45 | Pro 30 Glu | 15 Val Ala | Ala |
| 4 5 | Tyr As 1 Ile Hi Leu Pr | (()) SE p Gly s Pro e Phe 35 u Asn | A) NA B) LO QUENO Ala Met 20 Cys | AME/FOCATION Arg S Lys Leu Pro | ESCRI Leu Thr Leu Phe Arg | Val Ile Gly Glu 55 Glu | Tyr Val Ala 40 Leu | SEQ 1 Thr Arg 25 Val Ser | Leu 10 Tyr His Asp | D:343 Phe Ser Leu Gln Gln | Arg Arg Ser Ile | Asn Leu Val 45 Val | Pro 30 Glu Leu | 15 Val Ala Ser | Ala Gln Pro |
| 4 5 | Tyr As 1 Ile Hi Leu Pr Met Le | (()) SE Pro S Pro S Pro S Pro S Pro U Asn U Arg | A) NA B) LC QUENC Ala Met 20 Cys Thr | AME/FOCATION Arg Lys Leu Pro Tyr Phe | ESCRI Leu Thr Leu Phe Arg 70 | Val Ile Gly Glu 55 Glu | Tyr Val Ala 40 Leu | EEQ 1 Thr Arg 25 Val Ser Cys | Leu 10 Tyr His Asp Val | Phe Ser Leu Gln 75 | Arg Arg Ser Ile 60 Thr | Asn Leu Val 45 Val Lys | Pro 30 Glu Leu Glu | 15 Val Ala Ser Lys Ala | Ala Gln Pro Arg |
| 4 5 | Tyr As 1 Ile Hi Leu Ph Het Le 50 Thr G1 65 | (()) SE p Gly s Pro e Phe 35 u Asn u Arg | A) N/B) LC QUENC Ala Met 20 Cys Thr Gln Leu Gly | AME/FOCATION Arg Lys Leu Pro Tyr Phe 85 | Leu Thr Leu Phe Arg 70 Pro | Val Ile Gly Glu 55 Glu Leu | Tyr Val Ala 40 Leu Ile Ser | Thr Arg 25 Val Ser Cys Asp | Leu 10 Tyr His Asp Val Lys 90 | Phe Ser Leu Gln Gln 75 Leu | Arg Arg Ser Ile 60 Thr | Asn Leu Val 45 Val Lys Asp | Pro 30 Glu Leu Glu Ser | 15 Val Ala Ser Lys Ala 95 | Ala Gln Pro Arg 80 Tyr |
| 4 5 | Tyr As I le Hi Leu Pr Met Le 50 Thr Gl 65 Gly Al | (((()) SE Pro S Pro S Pro U Asn U Arg B Phe n Gly | A) N/B) LC QUENC Ala Het 20 Cys Thr Gln Leu Gly 100 | AME/H DCATI CE DE Arg Lys Leu Pro Tyr Phe 85 Ser | Leu Thr Leu Phe Arg 70 Pro | PTIC Val Ile Gly Glu 55 Glu Leu | Tyr Val Ala 40 Leu Ile Ser | Thr Arg 25 Val Ser Cys Asp Asp | Leu 10 Tyr His Asp Val Lys 90 Ile | Phe Ser Leu Gln 75 Leu | Arg Arg Ser Ile 60 Thr Arg | Asn Leu Val 45 Val Lys Asp | Pro 30 Glu Leu Glu Ser Tyr 110 | 15 Val Ala Ser Lys Ala 95 Leu | Ala Gln Pro Arg 80 Tyr |
| 45 50 55 | Tyr As Tyr As Leu Pr Met Le 50 Thr G1 65 Gly Al Val Ar | (()) SE p Gly s Pro e Fhe 35 u Asn u Arg a Asp g Phe n Gly 115 | A) N/B) LO QUENC Ala Met 20 Cys Thr Gln Leu Gly 100 Asn | AME/FOCATION Arg 5 Lys Leu Pro Tyr Phe 85 Ser Asn | Leu Thr Leu Phe Arg 70 Pro Ala | Val Ile Gly Glu 55 Glu Leu Tyr | Tyr Val Ala 40 Leu Ile Ser Gly Ser 120 | Thr Arg 25 Val Sec Cys Asp Asp Leu | Leu 10 Tyr His Asp Val Lys 90 Ile | Phe Ser Leu Gln Gln 75 Leu Ala | Arg Arg Ser Ile 60 Thr Arg Gly | Asn Leu Val 45 Val Lys Asp Asp Ser 125 | Pro 30 Glu Leu Glu Ser Tyr 110 Gly | 15 Val Ala Ser Lys Ala 95 Leu Gly | Ala Gln Pro Arg 80 Tyr Pro Arg |
| 45 50 55 60 | Tyr As 1 Ile Hi Leu Pr Met Le CThr Gl 65 Gly Al Val Ar Tyr As Ile Se | (()) SE p Gly s Pro e Phe 35 u Asn u Arg a Asp g Phe n Gly 115 0 | A) NA B) LO QUENO Ala Met 20 Cys Thr Gln Leu Gly 100 Asn | ANIE/FOCATION Arg S Lys Leu Pro Tyr Phe 85 Ser Asn Asn | Leu Thr Leu Phe Arg 70 Pro Ala Tyr | Val Ile Gly Glu 55 Glu Leu Tyr Ser Gly 135 | Tyr Val Ala 40 Leu Ile Ser Gly Ser 120 Thr | Thr Arg 25 Val Ser Cys Asp 105 Leu Leu | Leu 10 NC Leu 110 Tyr His Asp Val Lys 90 Ile Ser Gln | Phe Ser Leu Gln 75 Leu Ala Leu Gly | Arg Arg Ser Ile 60 Thr Arg Gly Glu Ser 140 | Asn Leu Val 45 Val Lys Asp Asp Ser 125 Ala | Pro 30 Glu Leu Glu Ser Tyr 110 Gly Ser | 15 Val Ala Ser Lys Ala 95 Leu Gly | Ala Gln Pro Arg 80 Tyr Pro Arg |
| 45 50 55 | Tyr As 1 Leu Pr Met Le 50 Thr Gl 65 Gly Al Val Ar Tyr As 11e Se 13 Arg Gl 145 | (()) SE p Gly s Pro e Fhe 35 u Asn u Arg a Asp g Phe n Gly 115 r Val 0 y Met | A) NA B) LO QUENO Ala Met 20 Cys Thr Gln Leu Gly 100 Asn Arg | ANIE/H Arg 5 Lys Leu Pro Tyr Phe 85 Ser Asn Lys | Leu Thr Leu Phe Arg 70 Pro Ala Tyr Arg 150 | Val Ile Gly Glu 55 Glu Leu Tyr Ser Gly 135 Ile | Tyr Val Ala 40 Leu Ile Ser Gly Ser 120 Thr | Thr Arg 25 Val Ser Cys Asp 105 Leu Leu Trp | Leu 10 Tyr His Asp Val Lys 90 Ile Ser Gln Asn | Phe Ser Leu Gln Gln 75 Leu Ala Leu Gly | Arg Arg Ser Ile 60 Thr Arg Gly Glu Ser 140 Leu | Asn Leu Val 45 Val Lys Asp Asp Ser 125 Ala | Pro 30 Glu Leu Glu Ser Tyr 110 Gly Ser Asn | 15 Val Ala Ser Lys Ala 95 Leu Gly Tyr | Ala Gln Pro Arg 80 Tyr Pro Arg Ser Glu 160 |
| 45 50 55 60 | Tyr As 1 1le Hi Leu Pr Met Le 50 Thr G1 65 Gly Al Val Ar Tyr As 11e Se 13 Arg Gl 145 Ala Ty | (()) SE p Gly s Pro e Phe 35 u Asn u Arg a Asp g Phe n Gly 115 r Val 0 y Met | A) NA B) LO QUENO Ala Met 20 Cys Thr Gln Leu Gly 100 Asn Arg His | AME/HE/HOCATI CE DE Arg 5 Lys Leu Pro Tyr Phe 85 Ser Asn Lys Tyr 165 | Leu Thr Leu Phe Arg 70 Pro Ala Tyr Tyr Arg 150 Leu | Val Ile Gly Glu 55 Glu Leu Tyr Ser Gly 135 Ile Val | Tyr Val Ala 40 Leu Ile Ser Gly Ser 120 Thr | Thr Arg 25 Val Ser Cys Asp 105 Leu Leu Trp | Leu 10 Tyr His Asp Val Lys Gln Asn Ser 170 | Phe Ser Leu Gln 75 Leu Ala Leu Gly Ala 155 | Arg Arg Ser Ile 60 Thr Arg Glu Ser 140 Leu Gly | Asn Leu Val 45 Val Lys Asp Asp Ser 125 Ala Arg | Pro 30 Glu Leu Glu Ser Tyr 110 Gly Ser Asn | 15 Val Ala Ser Lys Ala 95 Leu Gly Tyr Ala Tyr | Ala Gln Pro Arg 80 Tyr Pro Arg Ser Glu 160 His |
| 45 50 55 60 65 | Tyr As 1 Ile Hi Leu Ph Met Le 50 Thr G1 65 Gly Al Val Ar Tyr As Ile Se 145 Ala Ty Phe G1 | (()) SE p Gly s Pro e Phe 35 u Asn u Arg a Asp g Phe n Gly 115 r Val 0 Y Met r Tyr u Asp | A) NA B) LO QUENC Ala Met 20 Cys Thr Gln Leu Gly 100 Asn Arg His Pro | AME/HOCATI Arg 5 Lys Leu Pro Tyr Phe 85 Ser Asn Lys Lys Asn Lys | Leu Thr Leu Phe Arg 70 Pro Ala Tyr Arg 150 Leu Leu | Val Ile Gly Glu Leu Tyr Ser Gly 135 Ile Val | Tyr Val Ala 40 Leu Ile Ser Gly Ser 120 Thr Gly Ser Gly Ser | Thr Arg 25 Val Ser Cys Asp 105 Leu Trp Asp Tyr 185 | Leu 10 Tyr His Asp Val Lys 90 Ile Ser Gln Asn Ser 170 Tyr | Phe Ser Leu Gln 75 Leu Ala Leu Gly Ala 155 Thr | Arg Arg Ser Ile 60 Thr Arg Glu Ser 140 Leu Gly | Asn Leu Val 45 Val Lys Asp Asp Ser 125 Ala Arg Gly | Pro 30 Glu Leu Ser Tyr 110 Gly Ser Asn Asp | 15 Val Ala Ser Lys Ala 95 Leu Tyr Ala Tyr ₁₇₅ Gly | Ala Gln Pro Arg 80 Tyr Pro Arg Ser Glu 160 His |
| 45 50 55 60 | Tyr As 1 Leu Pr Met Le 50 Thr G1 65 Gly Al Val Ar Tyr As 11e Se Arg G1 145 Ala Ty Phe G1 Leu Pr | ((())) SE p Gly s Pro e Fhe 35 u Asn u Arg a Asp g Phe n Gly 115 r Val 0 y Met r Tyr u Asp o Leu 195 | A) NA B) LO QUENC Ala Met 20 Cys Thr Gln Leu Gly 100 Asn Arg His Pro Tyr 180 Gly | AME/HE/HOCATI Arg 5 Lys Leu Pro Tyr Phe 85 Ser Asn Lys Tyr 165 Arg Ile | Leu Thr Leu Phe Arg 70 Pro Ala Tyr Tyr Arg 150 Leu Leu Gly | Val Ile Gly Glu 55 Glu Leu Tyr Ser Gly 135 Ile Val Ala Phe | Tyr Val Ala 40 Leu Ile Ser Gly Ser Gly Ser Gly Ser Gly Ser Cly Ser | Thr Arg 25 Val Ser Asp 105 Leu Leu Trp Asp Tyr 185 Tyr | Leu 10 Tyr His Asp Val Lys 90 Ile Ser Gln Asn Ser 170 Tyr Arg | Phe Ser Leu Gln 75 Leu Ala Leu Gly Ala 155 Thr Ser Gly | Arg Arg Ser Ile 60 Thr Arg Glu Ser 140 Leu Gly Phe | Asn Leu Val 45 Val Lys Asp Asp Asp Gly Arg Gly Val 205 | Pro 30 Glu Leu Glu Ser Tyr 110 Gly Ser Asn Asp Ala | 15 Val Ala Ser Lys Ala 95 Leu Gly Tyr Ala Tyr ₁₇₅ Gly Tyr | Ala Gln Pro Arg 80 Tyr Pro Arg Ser Glu 160 His Arg |
| 45 50 55 60 65 | Tyr As 1 Ile Hi Leu Ph Met Le 50 Thr G1 65 Gly Al Val Ar Tyr As Ile Se 145 Ala Ty Phe G1 | (() (i) SE p Gly s Pro e Phe 35 u Asn u Arg a Asp g Phe 115 r Val y Met r Tyr u Asp o Leu 195 r Asp | A) NA B) LO QUENC Ala Met 20 Cys Thr Gln Leu Gly 100 Asn Arg His Pro Tyr 180 Gly | AME/HE/HOCATI Arg 5 Lys Leu Pro Tyr Phe 85 Ser Asn Lys Tyr 165 Arg Ile | Leu Thr Leu Phe Arg 70 Pro Ala Tyr Tyr Arg 150 Leu Leu Gly | Val Ile Gly Glu 55 Glu Leu Tyr Ser Gly 135 Ile Val Ala Phe | Tyr Val Ala 40 Leu Ile Ser Gly Ser Gly Ser Gly Ser Gly Ser Cly Ser | Thr Arg 25 Val Ser Asp 105 Leu Leu Trp Asp Tyr 185 Tyr | Leu 10 Tyr His Asp Val Lys 90 Ile Ser Gln Asn Ser 170 Tyr Arg | Phe Ser Leu Gln 75 Leu Ala Leu Gly Ala 155 Thr Ser Gly | Arg Arg Ser Ile 60 Thr Arg Glu Ser 140 Leu Gly Phe | Asn Leu Val 45 Val Lys Asp Asp Asp Gly Arg Gly Val 205 | Pro 30 Glu Leu Glu Ser Tyr 110 Gly Ser Asn Asp Ala | 15 Val Ala Ser Lys Ala 95 Leu Gly Tyr Ala Tyr ₁₇₅ Gly Tyr | Ala Gln Pro Arg 80 Tyr Pro Arg Ser Glu 160 His Arg |

| | Ala | Ala | туr | Leu | Tyr 245 | His | Arg | Gln | His | Leu 250 | Thr | Gln | Tyr | Asn | Trp 255 | Arg |
|----|------------|-------------|-------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------------|
| | Pro | Gĵλ | Gln | Gln 260 | | гла | Phe | Phe | Val 265 | | Туг | Gly | Phe | G1ÿ 270 | | Val |
| 5 | Aεp | Val | Ser 275 | | Ser | Pro | Ile | Trp 280 | | Gly | Ile | Ser | Arg 285 | _ | Asn | Tyr |
| | Val | Asn 290 | Gly | Trp | Lys | Leu | Ser 295 | Ser | Arg | Leu | Asp | Thr 300 | Arg | Arg | Gly | Asp |
| 10 | 305 | | G1 y | | | 310 | | | | | 315 | | | | | 320 |
| | | | Ile | | 325 | | | | | 330 | | | | | 335 | |
| 15 | | | Trp | 340 | | | | | 345 | | | | | 350 | | |
| 10 | | | Arg 355 Tyr | | | | | 360 | | - | | | 365 | | | |
| | | 370 | Leu | | | | 375 | | | | | 380 | | | | |
| 20 | 385 | | Asp | | | 390 | | | | | 395 | | | _ | - | 400 |
| | | | Туг | | 405 | | | | _ | 410 | | | _ | | 415 | |
| 25 | | | Leu | 420 | | | | | 425 | | _ | | | 430 | | |
| | Lys | | 435 Asp | Phe | Gly | Leu | | 440 Leu | Ser | Ala | Ala | | 445 Arg | Met | Val | Leu |
| 30 | | 450 His | Ser | Tyr | Ьуε | | 455 Arg | Thr | Ile | Gln | | 460 Glu | Gln | Leu | Asp | |
| 30 | 465 Gln | Leu | Ala | Tyr | | 470 Pro | туг | Ala | Туг | | 475 Asn | Arg | Glu | GJ?. | | 480 Glu |
| | Val | Arg | Ser | Ser 500 | 485 Leu | Tyr | Val | Ser | Ile 505 | 490 Pro | Met | Gln | Asn | Thr 510 | 495 His | Arg |
| 35 | Leu | Net | Thr 515 | | Leu | Arg | Leu | Tyr 520 | | Asp | Leu | Met | Lys 525 | | Lys | Asp |
| | Gly | 11e 530 | Ala | Tyr | G1 y | Lys | Thr 535 | | Gly | Val | Ile | Ser 540 | | Ile | Leu | Ser |
| 40 | Asp 545 | Pro | Gln | Ala | Glu | Arg 550 | Thr | Ser | Glγ | His | Thr 555 | Ile | Glу | Ala | lle | Су <i>s</i> 560 |
| | Asn | Ile | Ser | Tyr | Leu 565 | Phe | | | | | | | | | | |
| 45 | (2) | INEC | ORHAT | HOI | FOR | SEQ | ID 1 | 10:34 | 14 | | | | | | | |
| 40 | | (i) | SEÇ | UENC | | | | | | 1 | | | | | | |
| | | | (E | 3) Ti | PE: | amir | o ac | cid | acic | 15 | | | | | | |
| 50 | | (ii) | HOL | | | | | | | | | | | | | |
| | , | (iii) | HYE | OTHE | TICE | L: Y | ŒS | | | | | | | | | |
| 55 | | (vi) | ORI | | | | | | | | | | | | | |
| | | | | | | SM: | Porp | hyro | mona | as gi | ingiv | /alis | ? | | | |
| 60 | | (1X) | | i) HP | 1-1E/F | | | | ture | <u> </u> | | | | | | |
| 00 | | 1211 | SEÇ |) Lo | | | | | EC I | וים אור | 7.34 | 1 | | | | |
| | Ara | | Asn | | | | | | | | | | Glv | Ara | Ara | Glv |
| 65 | 1 | | Сув | | 5 | | | | | 10 | _ | | _ | _ | 15 | _ |
| | | | Glu | 20 | | | | | 25 | | | | | 30 | | |
| 70 | | Hi <i>s</i> | 35 Gly | | | | Ile | 40 | | _ | | | 45 | | | _ |
| | | 50 Asp | Lys | Arg | Phe | | 55 Lys | Lys | Leu | His | | Ile | Ala | Ala | Leu | |
| 75 | 65 Val | Leu | Pro | Phe | | 70 Leu | Thr | Ala | Gln | | 75 Pro | Val | Ser | Asn | | 80 Glu |
| 75 | | | | | 85 | | | | | 90 | | | | | 95 | |

| | Ile | Asp | Ser | Leu 100 | Ser | Asn | Val | Gln | Leu 105 | | Thr | Val | Gln | Val 110 | | Ala |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------------------|------------|------------|------------|------------|------------|
| | Thr | Arg | Ala 115 | Thr | Ala | Lys | Thr | Pro 120 | | Ala | Tyr | Thr | Asn 125 | Val | Arg | Lys |
| 5 | Ala | Glu 130 | Leu | Ser | Lys | Ser | Asn 135 | | Glγ | Arg | Asp | Ile 140 | Pro | Туг | Leu | Leu |
| | Het 145 | Leu | Thr | Pro | Ser | Val 150 | | Ala | Thr | Ser | Asp 155 | Ala | | Thr | Gly | 11e |
| 10 | G1?. | Tyr | Ser | Gly | Phe 165 | Arg | Val | Arg | Gly | Thr 170 | | | Asn | Arg | Ile 175 | Asn |
| | Ile | Thr | Thr | Asn 180 | Gly | Val | Pro | Leu | Asn 185 | | Ser | Glu | Ser | Gln 190 | Ser | Val |
| | Phe | Trp | Val 195 | Asn | Met | Pro | Asp | Phe 200 | | Ser | Ser | Ile | Glu 205 | | Leu | Gln |
| 15 | Val | Gln 210 | Arg | Gly | Vāl | Gly | Thr 215 | | Thr | Asn | Gly | Ala 220 | | Ala | Phe | Gly |
| | Ala 225 | Ser | Val | Asn | Met | Arg 230 | Thr | Asp | Asn | Leu | Gl ₃ . 235 | | Ala | Pro | Tyr | Gly 240 |
| 20 | Arg | Val | Asp | Leu | Ser 245 | Glγ | Gly | Ser | Phe | Gly 250 | Thr | Phe | Arg | Arg | Ser 255 | Val |
| | Γ?.ε | Leu | Gly | Ser 260 | Glу | Arg | Ile | Gly | Arg 265 | | | Ala | Val | Asp 270 | Ala | Arg |
| | Leu | Ser | Lys 275 | Ile | Gl y | Ser | Asp | Gly 280 | Tyr | Vāl | Asp | Arg | Gly 285 | Ser | Val | Asp |
| 25 | | 290 | | Tyr | | | 295 | | | | | 300 | | | | |
| | 305 | | | Ile | | 310 | | | | | 315 | | | | | 320 |
| 30 | | | | Ser | 325 | | | | | 330 | | | | | 335 | |
| | | | | Leu 340 | | | | | 345 | | | | | 350 | | |
| | | | 355 | Asp | | | | 360 | | | | | 365 | | | |
| 35 | | 370 | | Ser | | | 375 | | | | | 380 | | | | |
| | 385 | | | Gly | | 390 | | | | | 395 | | | | | 400 |
| 4 0 | | | | Leu | 405 | | | | | 410 | | | | | 415 | |
| | | | | 11e 420 | | | | | 425 | | | | | 430 | | |
| 45 | | | 435 | Leu | | | | 440 | | | | | 445 | | | |
| 40 | | 450 | | Asn | | | 455 | | | | | 460 | | | | |
| | 465 | | | Tyr | | 470 | | | | | 475 | | | | | 480 |
| 50 | | | | Asp | 485 | | | | | 490 | | | | | 495 | |
| | | | | Pro 500 | | | | | 505 | | | | | 510 | | |
| 55 | | | 515 | Thr Gln | | | | 520 | | | | | 525 | | | |
| | | 530 | | | | | 535 | | | | | 540 | | | | |
| | 545 | | | Gly Ala | | 550 | | | | | 555 | | | | | 560 |
| 60 | | | | Ile | 565 | | | | | 570 | | | | | 575 | |
| | | | | 580 Tyr | | | | | 585 | | | | | 590 | | |
| 65 | | | 595 | Net | | | | 600 | | | | | 605 | | | |
| | | 610 | | Gly | | | 615 | | | | | 620 | | | | |
| | 625 | | | Glu | | 630 | | | | | 635 | | | | | 640 |
| 70 | | | | Ala | 645 | | | | | 650 | | | | | 655 | |
| | | | | 660 Thr | | | | | 665 | | | | | 670 | | |
| <i>7</i> 5 | | | 675 | Thr | | | | 680 | | | | | 685 | | | |
| | J | | | | | | | 1 | p | 115 | , 11 a | . y L | 261 | | Well. | Val |

```
700
       Ile Ala Gly Ser Met Leu Thr Leu Ser His Ala Gly Phe Glu Met Ala 705 710 715 720
       Trp Thr Ser Arg Phe Val Ser Lys Gln Tyr Leu Asp Asn Thr Gln Arg
725 730 735
 5
       Ser Asp Arg Met Leu Ser Ser Tyr Trp Val Asn Asp Leu Arg Leu Gly 740 745 750
       Tyr Val Leu Pro Val His Phe Val Lys Arg Val Ala Leu Gly Val Gln
755 760 765
       Leu Asn Asn Leu Phe Asn Leu Met Tyr Ala Ser Asn Ala Tyr Ile Tyr 770 775 780
10
       Asp Ala Gly Tyr Val Gln Ala Ser Gly Glu Leu Ser Ala Tyr Ala Asp 785 790 795 800
                                                    795
       Leu Arg Tyr Tyr Pro Gln Ala Gly Phe Asn Ala Leu Gly Ser Leu Thr
15
       Ile Asp Phe
       (2) INFORMATION FOR SEQ ID NO: 345
20
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 532 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
30
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
35
                  (B) LOCATION 1...532
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:345
       Tyr Arg Arg Ser Gly Arg Val Cys Pro Arg Ile Leu Arg Asn Lys Arg
40
                                               10
       Ser Tyr Ala Asp Thr Arg Leu Gln Pro Arg Arg Ala Val Arg Leu Arg
20 25 30
       Ala Gly Thr Arg Thr Lys Met Lys Arg Arg Phe Leu Ser Leu Leu 35 40 45
       Leu Tyr Ile Leu Ser Ser Ile Ser Leu Ser Ala Gln Arg Phe Pro Met
50 60
45
       Val Gln Gly Ile Glu Leu Asp Thr Asp Ser Leu Phe Ser Leu Pro Lys
65 70 75 80
       Arg Pro Trp Arg Ala Ile Gly Lys Thr Ile Gly Val Asn Leu Ala Val
85 90 95
50
       Trp Gly Phe Asp His Phe Ile Met Asn Glu Asp Phe Ala Asp Ile Ser
100 105 110
       Trp Gln Thr Ile Lys Ser Asn Phe Gln Thr Gly Phe Gly Trp Asp Asn 115 120 125
55
       Asp Lys Phe Val Thr Asn Leu Phe Ala His Pro Tyr His Gly Ser Leu
130 135 140
       Tyr Phe Asn Ala Ala Arg Ser Asn Gly Leu Ser Phe Arg His Ser Ala
145 150 155 160
       Pro Phe Ala Phe Phe Gly Ser Leu Met Trp Glu Leu Leu Met Glu Asn 165 170 175
60
       Glu Pro Pro Ser Ile Asn Asp Leu Cys Ala Thr Thr Ile Gly Gly Ile
180 185 190
       Ala Leu Gly Glu Met Gly His Arg Leu Ser Asp Leu Leu Ile Asp Asn
195 200 205
65
       Arg Thr Thr Gly Trp Glu Arg Met Gly Arg Glu Val Ala Ile Ala Leu
210 215 220
       Ile Asn Pro Met Arg Phe Leu Asn Arg Leu Thr Ala Gly Glu Val Thr 225 230 235 240
       Ser Val Gly Ser Arg Ser Gly Gln Ile Phe Gln Ser Val Pro Ile Asn
245 250 255
70
       Ile Val Val Asp Ala Gly Phe Arg Phe Leu Ala Asp Lys Arg His Ala 260 265 270
       Arg Thr Gly Ala Thr Ala Leu Thr Leu Asn Leu Arg Phe Asp Tyr Gly 275 280 285
                                      280
75
       Asp Pro Phe Arg Ser Glu Thr Phe Ser Pro Tyr Asp Phe Phe Gln Phe
```

```
Lys Ala Gly Leu Ser Phe Ser Glu Ser Gln Pro Leu Leu Ser Gln Ile
305 310 315 320
        Asn Leu Ile Gly Ile Leu Ser Gly Cys Gln Leu Leu Ala His Glu Arg
325 330 335
  5
       Thr Val Leu Val Gly Gly Leu Phe Gln His Phe Asp Tyr Tyr Asn Ser 340 345 350
        Glu Lys Arg Ile Ser Lys Asn Ser Glu Glu Val Leu Val Thr Pro Tyr
355 360 365
       Arg Ile Ser Gln Val Ala Ala Leu Gly Gly Gly Leu Ile Phe Gln His
370 375 380
10
       His Gly Lys Phe Arg Arg Arg Pro Leu Glu Leu Tyr Ala Glu Thr Tyr
385 390 395 400
       Leu Asn Val Val Pro Met Gly Ala Ser Leu Ser Asp His Tyr Asn Val
405 410 415
15
       Asp Asn Arg Asp Tyr Asn Leu Gly Ser Gly Leu Ser Gly Lys Leu Tyr 420 425 430
       Leu Gly Ala Thr Tyr Asn Asp Leu Trp Ser Trp Leu Leu Gly Val Glu
435 440 445
20
       Ser Tyr Arg Leu Tyr Thr Trp Ile Gly Tyr Glu Glu Pro His Gln Lys
450 455 460
       Asn Thr Asp Val Ser Ser Phe Met Val Gln Gly Asp Glu Ser Lys Ala
465 470 475 480
       Arg Leu Leu Val Thr Ser Ser Glu Phe Ala Phe His Pro Gly Pro Trp
485 490 495
25
       His Val Ala Ile Val Ala Arg Arg Phe Ile Arg Lys Thr Ala Tyr Gln 500 505 510
       Phe Tyr Pro Asn Val Ser Phe Asp Thr Gly Asp Ile Gln Leu Arg Val 515 525 525
30
       Gly Phe His Phe
            530
        (2) INFORMATION FOR SEQ ID NO:346
35
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 300 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
40
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
45
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                  (B) LOCATION 1...300
50
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:346
       Lys Arg Arg Lys Lys Gln Met Lys Arg Leu Ile Val Phe Leu Ala Met
                                               10
55
       Gly Gly Leu Leu Phe Thr Leu Ala Asn Ala Gln Glu Ala Asn Thr Ala 20 25 30
       Ser Asp Thr Pro Lys Lys Asp Trp Thr Ile Lys Gly Val Thr Gly Leu 35 40
       Asn Ala Ser Gln Thr Ser Leu Thr Asn Trp Ala Ala Gly Gly Glu Asn 50 60
60
       Thr Val Ala Gly Asn Leu Tyr Leu Asn Ile Asp Ala Asn Tyr Leu Lys
65 70 75 80
       Asp Lys Trp Ser Trp Asp Asn Gly Leu Arg Thr Asp Phe Gly Leu Thr
85 90 95
65
       Tyr Thr Thr Ala Asn Lys Trp Asn Lys Ser Val Asp Lys Ile Glu Leu
100 105 110
       Fhe Thr Lys Ala Gly Tyr Glu Ile Gly Lys His Trp Tyr Gly Ser Ala 115 120 125
       Leu Phe Thr Phe Leu Ser Gln Tyr Ala Lys Gly Tyr Glu Lys Pro Ser
130 135 140
70
       Asp His Leu Thr Gly Val Lys His Ile Ser Asn Phe Phe Ala Pro Ala
145 150 155 160
       Tyr Leu Thr Leu Gly Ile Gly Ala Asp Tyr Lys Pro Asn Glu Lys Phe
165 170
75
       Ser Leu Tyr Leu Ser Pro Thr Thr Gly Lys Leu Thr Val Val Ala Asp
```

```
180
                                       185
      Asp Tyr Leu Ser Ser Leu Gly Ala Phe Gly Val Lys Val Gly Glu Lys
             195
                                  200
      Thr Met Phe Glu Leu Gly Ala Leu Val Val Gly Ser Ala Asn Ile Asn 210 215 220
 5
                              215
         210
      Leu Met Glu Asn Val Asn Leu Ile Thr Lys Ala Ser Phe Phe Ser Ala
                         230
                                              235
      225
      Tyr Thr His Asp Phe Gly Asn Ile Asp Ile Asn Trp Glu Ala Met Leu
                                          250
                      245
      Ala Met Lys Ile Asn Lys Phe Leu Thr Ala Thr Ile Ala Thr Asn Leu
260 265 270
10
      Ile Tyr Asp Asp Asp Val Lys Ile Asn Asp Gly Pro Lys Ile Gln Phe 275 280 285
      Lys Glu Val Val Gly Val Gly Val Ala Tyr Thr Phe
15
                              295
      (2) INFORMATION FOR SEO ID NO:347
           (i) SEQUENCE CHARACTERISTICS:
20
                (A) LENGTH: 221 amino acids
                (B) TYPE: amino acid
                (D) TOPOLOGY: linear
          (ii) MOLECULE TYPE: protein
25
         (iii) HYPOTHETICAL: YES
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: Porphyromonas gingivalis
30
          (ix) FEATURE:
                (A) NAME/KEY: misc feature
                (B) LOCATION 1...221
35
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 347
      Thr Arg Glu Ser Val Leu His Cys Arg Thr Lys Leu Lys Lys Glu Arg
                                           10
      Lys Met Lys Lys Met Ile Leu Ala Ala Thr Met Leu Leu Ala Thr Ile
40
                  20
                                      25
                                                          30
      Gly Phe Ala Asn Ala Gln Ser Arg Pro Ala Leu Arg Leu Asp Ala Asn
                                  40
      Phe Val Gly Ser Asn Leu Met Gln Lys Val Ala Asn Thr Ser Val Asn
50 60
45
      Asn Lys Met Ile Val Gly Leu Arg Val Gly Ala Ala Ala Glu Phe Ala
65 75 80
      Leu Ser Asn Asp Gly Phe Tyr Leu Ala Pro Gly Leu Ala Tyr Thr Met
                     85
                                          90
      Arg Gly Ala Lys Met Glu Ser Leu Ser Glu Thr Thr Arg Leu His
50
                 100
                                     105
                                                          110
      Tyr Leu Gln Ile Pro Val Asn Ala Gly Met Arg Phe Ser Phe Ala Asp
             115
                                120
      Asn Met Ala Ile Ser Leu Glu Ala Gly Pro Tyr Phe Ala Tyr Gly Val
        130
                            135
                                                  140
55
      Ala Gly Thr Ile Lys Thr Lys Val Ala Gly Val Thr Ala Ser Val Asp 150 155 160
      Ala Phe Gly Asp Asn Gly Tyr Asn Arg Phe Asp Leu Gly Leu 165 170 175
      Ser Ala Ala Leu Ser Tyr Asp Arg Tyr Tyr Val Gln Ile Gly Tyr Glu
180 185 190
60
                                   185
      His Gly Leu Leu Asn Met Leu Lys Asp Ala Pro Asp Lys Thr Ser Leu
             195 200
      Arg Asn His Asp Phe Phe Val Gly Leu Gly Val Arg Phe
                            215
65
      (2) INFORMATION FOR SEQ ID NO:348
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 248 amino acids
70
                (B) TYPE: amino acid
                (D) TOPOLOGY: linear
          (ii) MOLECULE TYPE: protein
75
         (iii) HYPOTHETICAL: YES
```

```
(vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
 5
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...248
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:348
10
       Ile Lys Arg Ile Glu Met Lys Arg Ile Phe Thr Val Ala Leu Val Leu 1 5 10 15
       Leu Ala Ser Val Thr Met Ala Ile Gly Gln Ser Arg Pro Ala Leu Arg
20 25 30
       Val Asp Ala Asn Phe Val Gly Ser Asn Gln Ser Met Lys Arg Asp Gly 35 40 45
       Tyr Val Trp Asp Thr Lys Met Asn Val Gly Leu Arg Val Gly Ala Ala 50 55 60
       Ala Glu Phe Met Ile Gly Ser Arg Gly Phe Tyr Leu Ala Pro Gly Leu
65 70 75 80
20
      Asn Tyr Thr Het Lys Gly Ser Lys Thr Glu Trp Asp Ile Pro Glu Met 85 90 95
       Val Pro Gly Thr Tyr Ile Thr Met Val Ser Thr Arg Leu His Tyr Leu
100 105 110
25
       Gln Leu Pro Ile Asn Ala Gly Net Arg Phe Asp Leu Het Asn Asp Met 115 120 125
      Ala Val Ser Ile Glu Ala Gly Pro Phe Leu Ala Tyr Gly Ile Tyr Gly
130 135 140
       Thr Tyr Arg Gln Lys Leu Glu Gly Trp Lys Pro Asn Asn Tyr Ser Thr 145 150 160
30
       Glu Phe Phe Gly Pro Thr Leu Gly Gly Pro Thr Asn Ile Arg Trp Asp
165 170 175
       Ile Gly Ala Asn Ile Ile Ala Ala Phe His Tyr Lys Arg Tyr Tyr Ile 180 $180$
35
       Gln Ile Gly Tyr Glu His Gly Phe Val Asp Ile Val Ser Gly Gly 195 200 205
       Ser Asp Ile Pro Arg Leu Asn Asp Asn Arg Gln Ser Ser Ser Thr Thr
210 215 220
       Ala Leu Arg Glu Lys Gly Asn Asn Glu Tyr Ala Tyr Asn Arg Asp Phe
225 230 235 240
40
       Phe Val Gly Ile Gly Tyr Arg Phe
       (2) INFORMATION FOR SEQ ID NO:349
45
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 211 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
50
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
55
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
60
                  (B) LOCATION 1...211
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:349
      Lys Arg Lys Ser Met Lys Arg Met Leu Leu Leu Leu Val Val Leu Leu 1 5 10 15
65
      Tyr Gly Ile Ala Gly Arg Leu Ala Ala Gln Asp Val Ile Arg Pro Trp
20 25 30
      Ser Leu Gln Val Gly Ala Gly Tyr Ser Asp Thr Giu Asn lie Pro Gly 35 40 45
70
      Gly Phe Thr Tyr Gly Phe Tyr Leu Gly Lys Arg Het Gly Ser Phe Leu 50 60
       Glu Val Gly Leu Ser Met Tyr Asn Ser Thr Arg Gln Thr Ala Asn Asn 65 70 80
      Ala Asp Ser Phe Ala Ser Asn Glu Gly Asp Gly Ser Phe Gln Val Asn
85 90 95
75
```

```
Met Ser Ser Pro Asn Glu Lys Trp Ser Phe Phe Asp Ala Gly Ser Ala 100 105 110
       Asn Cys Tyr Net Ile Val Val Gly Val Asn Pro Leu His Leu Phe Trp
115 120 125
       Gln Asn Ser Arg His Asn Leu Phe Leu Ala Val Gln Ala Gly Leu Ser
130 135 140
       Asn Lys His Asn Ile His Phe Ile Tyr Gly Asp Lys Gly Ala Lys Val
145 150 155 160
       Ser Ile Tyr Thr Asn Ser Asn Thr Tyr Ile Gly Tyr Gly Ala Arg Val
165 170 175
10
       Ala Tyr Glu Tyr Gln Ile His Lys Asn Val Gly Ala Gly Ala Ala Val 180 185 190
       Met Tyr Asp His Gly Asn Lys Met Leu Thr Ala Met Ala Thr Leu Ser
              195
15
       Thr His Phe
           210
       (2) INFORMATION FOR SEQ ID NO:350
20
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 953 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
30
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...953
35
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:350
       Ile Arg Met Arg Val Ser Asp Leu Cys Ser Arg Leu Ser Trp Leu Leu 1 5 10 10 15
       Pro Val Ile Leu Val Gly Leu Leu Cys Ala Thr Leu Val Ala Ala Glu
20 25 30
40
       Arg Pro Met Ala Gly Ala Val Gly Leu His His Arg Arg His Ala Ala
35 40 45
       Leu Ser Asp Ser Thr Ala Lys Asp Thr Val Pro Leu Ala Lys Pro Ile
50 55 60
       Pro Asp Ser Ala Phe Arg Asp Ser Leu Pro Ala Asp Ser Thr Gly Ser
65 70 75 80
      Het Arg Gln Asp Ser Val Tyr Asp Asp Glu Phe Glu Leu Glu Asp Ile 85 90 95
       Val Glu Tyr Glu Ala Ala Asp Ser Ile Val Leu Leu Gly Gln Asn Arg
100 105 110
50
       Ala Tyr Leu Phe Gly Lys Ser Tyr Val Ser Tyr Gln Lys Ser Arg Leu
115 120 125
       Glu Ala Asn Phe Met Tyr Leu Asn Thr Asp Ser Ser Thr Val Tyr Thr
130 140
55
       Arg Tyr Val Leu Asp Thr Ala Gly Tyr Pro Net Ala Phe Pro Val Phe 145 150 155 160
       Lys Asp Gly Glu Gln Ser Phe Glu Ala Lys Asn Phe Thr Tyr Asn Phe
165 170 175
60
       Arg Thr Glu Lys Gly Ile Ile Ser Gly Val Ile Thr Gln Gin Gly Glu
180 185 190
       Gly Tyr Leu Thr Ala Gly Lys Thr Lys Lys Het Pro Asp Asn Ile Met
195 200 205
       Phe Met Gin Gly Gly Arg Tyr Thr Thr Cys Asp Asn His Asp His Pro
210 215 220
65
       His Phe Tyr Ile Asn Leu Ser Lys Ala Lys Val His Pro Glu Lys Asp
225 230 235 240
       Ile Val Thr Gly Pro Val Asn Leu Val Ile Ala Asp Met Pro Leu Pro
245 250 255
70
       Ile Gly Leu Pro Phe Gly Tyr Phe Pro Phe Ser Asn Lys Tyr Ser Ser 260 265 270
       Gly Ile Leu Met Pro Thr Tyr Gly Glu Asp Asn Arg Tyr Gly Phe Tyr
275 280 285
       Leu Arg Asn Gly Gly Tyr Tyr Phe Ala Phe Ser Asp Tyr Ile Asp Leu
290 295 300
75
```

| | Al a 305 | Leu | Arg | Gly | Glu | Ile 310 | Phe | Ser | Lys | Gly | Ser 315 | Trp | Gly | Ile | Ser | Ala 320 |
|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | | | Lys | Туг | Lys 325 | Lys | | Туг | Lys | Tyr 330 | Asn | | Ser | Phe | G1u 335 | Ala |
| 5 | Asn | Туг | Leu | Val 340 | Ser | | Ser | Gly | Asp 345 | Lys | | Val | Pro | Gly 350 | Asp | |
| | Ser | L).s | Thr 355 | Thr | Ser | Leu | Asn | 11e 360 | Arg | | Thr | His | Ser 365 | Gln | Asp | Pro |
| 10 | Γ7.ε | Ala 370 | Asn | Pro | Leu | Gln | Thr 375 | Leu | Ser | Ala | Asn | Val 380 | Asn | Phe | Ala | Thr |
| | 385 | | | Phe | | 390 | | | | | 395 | | | | | 400 |
| 45 | | | | Thr | 405 | | | | | 410 | | | | | 415 | |
| 15 | | | | Pro 420 | | | | | 425 | | | | | 430 | | |
| | | | 435 | | | | | 440 | | | | | 445 | | | |
| 20 | | 450 | | Arg | | | 455 | | | | | 460 | | | | |
| | 465 | | | Glu | | 470 | | | | | 475 | | | | | 480 |
| 25 | | | | Thr | 485 | | | | | 490 | | | | | 495 | |
| 23 | | | | Asn 500 | | | | | 505 | | | | | 510 | | |
| | | | 515 | | | | | 520 | | | | | 525 | | | |
| 30 | | 530 | | Thr Pro | | | 535 | | | | | 540 | | | | |
| | 545 | | | Ser | | 550 | | | | | 555 | | | | | 560 |
| 35 | | | | Pro | 565 | | | | | 570 | | | | | 575 | |
| | | | | 580 Pro | | | | | 585 | | | | | 590 | | |
| | | | 595 | Gly | | | | 600 | | | | | 605 | | | |
| 40 | | 610 | | Thr | | | 615 | | | | | 620 | | | | |
| | 625 | | | Net | | 630 | | | | | 635 | | | | | 640 |
| 45 | Asn | Leu | Glu | Ala | 645 Lys | Ile | Lys | Ser | Lys | | Asp | Ser | Thr | Gly | 655 Ile | Lys |
| | Lys | Ile | Ser | 660 Leu | Ile | Asp | Gln | Phe | 665 Thr | | Ser | Thr | | 670 Tyr | Asn | Met |
| 50 | Phe | Ala 690 | 675 Asp | Ser | Ile | Arg | Trp | 680 Ser | Asn | Ile | Ser | | 685 Ser | Leu | Alā | Leu |
| 00 | Arg | | Ser | Lys | Ser | Phe | 695 Thr | Leu | Arg | Leu | | 700 Gly | Leu | Phe | Asp | |
| | | Leu | Thr | Lys | Tyr 725 | | Glu | Gly | Glu | Asp 730 | 715 Gly | Lys | Ile | lle | | 720 Tyr |
| 55 | Lys | Ser | Asn | Asp 740 | Leu | Arg | Ile | Phe | Asn 745 | Gly | Lys | Gly | Leu | Ala 750 | 735 Arg | Leu |
| | Ile | Ser | Thr 755 | Gly | Thr | Ser | Phe | Ser 760 | Tyr | Thr | Leu | Asn | Lys 765 | Glu | Ser | Leu |
| 60 | Ser | Gly 770 | Leu | Ile | Ala | Leu | Phe 775 | | Gly. | Lys | Lys | Glu 780 | Arg | Arg | Asp | Glu |
| | Lys 785 | Lys | Asn | Thr | Gly | Ala 790 | Thr | Pro | His | Glu | Gly 795 | | Asp | Ala | Ala | Asp 800 |
| 0.5 | | | | Gly | 805 | | | | | 810 | Ser | | | | 815 | Leu |
| 65 | | | | Arg 820 | | | | | 825 | | | | | 830 | | |
| | | | 835 | Pro | | | | 840 | | | | | 845 | | | |
| 70 | | 850 | | Asn | | | 855 | | | | | 860 | | | | |
| | 865 | | | Asn | | 870 | | | | | 875 | | | | | 880 |
| 75 | | | | Gly | 885 | | | | | 890 | | | | | 895 | |
| <i>7</i> 5 | Thr | Ser | Leu | Thr | Cys | Asn | Val | Thr | Arg | Asp | llet | His | Суѕ | Trp | Ala | Ile |

```
905
       Ser Ala Ser Phe Ile Pro Ile Gly Ala Tyr Lys Ser Tyr Asn Phe Val
915 920 925
       Ile Ser Val Lys Ser Ser Leu Leu Gln Asp Leu Lys Tyr Gln Gln Ser
930 935 940
       Asn Arg Pro Ile Thr Asn Thr Trp Tyr
       945
        (2) INFORMATION FOR SEQ ID NO:351
10
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1251 amino acids
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
15
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
20
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAM4E/KEY: misc_feature
                   (B) LOCATION 1...1251
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:351
       30
       Phe Val Val Thr Asn Glu Val Ile Ile Met Met Lys Arg Tyr Thr Ile
20 25 30
       Ile Leu Ala Val Phe Leu Leu Phe Cys Thr Val Phe Thr Phe Gln Ile
35 40 45
35
       Lys Ala Arg Pro Tyr Glu Arg Phe Ala Asp Val Glu Lys Pro Trp Ile 50 \hspace{1cm} 55
       Gln Lys His Ser Met Asp Ser Lys Leu Val Pro Ala Asn Lys Gly Asn 65 70 75 80
       Leu Ile Gln Ala Glu Ile Val Tyr Gln Ser Val Ser Glu His Ser Asp
85 90 95
40
       Leu Val Ile Ser Pro Val Asn Glu Ile Arg Pro Ala Asn Arg Phe Pro
100 105 110
       Ser His Arg Lys Ser Phe Phe Ala Glu Asn Leu Arg Ala Ser Pro Pro
115 120 125
45
       Val Val Pro Val Ala Val Asp Lys Tyr Ala Val Pro Val Ala Asn Pro
130 135 140
       Het Asp Pro Glu Asn Pro Asn Ala Trp Asp Val Thr Leu Lys Ile Thr 145 150 155 160

Thr Lys Ala Val Thr Val Pro Val Asp Val Val Met Val Ile Asp Gln 165 170 170 175
50
       Ser Ser Ser Net Gly Gly Gln Asn Ile Ala Arg Leu Lys Ser Ala Ile
180 185 190
       Ala Ser Gly Gln Arg Phe Val Lys Lys Met Leu Pro Lys Gly Thr Ala
195 200 205
55
       Thr Glu Gly Val Arg Ile Ala Leu Val Ser Tyr Asp His Glu Pro His
210 215 220
       Arg Leu Ser Asp Phe Thr Lys Asp Thr Ala Phe Leu Cys Gln Lys Ile
225 230 235 240
       Arg Ala Leu Thr Pro Ile Trp Gly Thr His Thr Gln Gly Gly Leu Lys 245 250 255
60
       Het Ala Arg Asn Ile Het Ala Thr Ser Thr Ala Val Asp Lys His Ile
260 265 270
       Ile Leu Met Ser Asp Gly Leu Ala Thr Glu Gln Tyr Pro Val Lys Asn
275 280 285
65
       Val Thr Thr Ala Asp Phe Ile Gly Lys Thr Gly Asn Ala Asn Asp Pro
290 295 300
       Ile Asp Leu Val Ile Gln Gly Ala Ile Asn Phe Pro Thr Asn Tyr Val 305 310 315 320
       Ser Asn Asn Pro Ser Thr Pro Leu Thr Pro Asn Tyr Pro Thr His Ser 325 330 335
70
       Ser Lys Val Gly Arg Arg Asn Leu Pro Glu Ser Lys Phe Asp Tyr Ser 340 345 350
       Asn Leu Ser Ala Arg Ile Thr Phe Asp Gly Val Ala Gly Ala Leu Val
355 360 365
       Tyr Glu Pro Arg Phe Pro His Pro Tyr Tyr Tyr Phe Pro Cys Asn
```

| 370 Ala Ala 1le Asn Glu Ala Gln Phe Ala Lys Asn Ser Gly 385 385 386 His Thr Ile Gly Tyr Asp Leu Gly Asp Phe Ala Leu Ala 405 Leu Lys Leu Thr Ala Thr Asp Glu Asn His Phe Phe Thr 420 Ala Asn Leu Ala Ala Ala Phe Asp Asn Ile Ala Gln Thr 435 Gly Ile Gln Arg Gly Glu Val Thr Asp Phe Val Ala Pro 450 Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr His Leu 465 Ser Asn Gly Thr Val His Tyr Asp Val Ser Thr Lys Lys 485 Thr Thr Gly Thr Ile Leu Ser Ser Ser Glu Ala Thr Ile 500 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asn Asp Ile 515 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp 530 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp 540 | a Asn A r Ala T 430 r Ile A 5 o Gly E u Leu A s Leu T e Thr T 510 e Pro V 5 o Thr A | 400 Asn Ser 415 Thr Pro Asn Ile Phe Ile Asn Val 480 Thr Trp 495 Tyr Arg |
|---|---|--|
| 10 His Thr Ile Gly Tyr Asp Leu Gly Asp Phe Ala Leu Ala 405 Leu Lys Leu Thr Ala Thr Asp Glu Asn His Phe Phe Thr 420 Ala Asn Leu Ala Ala Ala Phe Asp Asn Ile Ala Gln Thr 435 Gly Ile Gln Arg Gly Glu Val Thr Asp Phe Val Ala Pro 455 Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr His Leu 465 Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr His Leu 465 Ser Asn Gly Thr Val His Tyr Asp Val Ser Thr Lys Lys 485 Thr Thr Gly Thr Ile Leu Ser Ser Ser Glu Ala Thr Ile 500 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asn Asp Ile 515 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | r Ala T 430 r Ile A 5 o Gly E u Leu A e Thr T 510 e Pro V 5 o Thr A | Asn Ser 415 Thr Pro Asn Ile Phe Ile Asn Val 480 Thr Trp 495 Tyr Arg |
| Leu Lys Leu Thr Ala Thr Asp Glu Asn His Phe Phe Thr 420 Ala Asn Leu Ala Ala Ala Phe Asp Asn Ile Ala Gln Thr 435 10 Gly Ile Gln Arg Gly Glu Val Thr Asp Phe Val Ala Pro 450 Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr His Leu 465 Ser Asn Gly Thr Val His Tyr Asp Val Ser Thr Lys Lys 485 Thr Thr Gly Thr Ile Leu Ser Ser Glu Ala Thr Ile 500 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asn Asp Ile 515 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | r Ala T 430 r Ile F 5 o Gly E u Leu A s Leu T e Thr T 510 e Pro V 5 o Thr A | Thr Pro Asn Ile Phe Ile Asn Val 480 Thr Trp 495 Tyr Arg |
| Ala Asn Leu Ala Ala Ala Phe Asp Asn Ile Ala Gln Thr 435 10 Gly Ile Gln Arg Gly Glu Val Thr Asp Phe Val Ala Pro 450 Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr His Leu 460 Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr Lys Lys 485 Thr Thr Gly Thr Val His Tyr Asp Val Ser Thr Lys Lys 485 Thr Thr Gly Thr Ile Leu Ser Ser Glu Ala Thr Ile 500 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asn Asp Ile 515 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | r Ile A 5 0 Gly E u Leu A 5 Leu T 4 e Thr T 510 e Pro V 5 p Thr A | Phe Ile Asn Val 480 Thr Trp 495 Tyr Arg Val Asn |
| 10 Gly Ile Gln Arg Gly Glu Val Thr Asp Phe Val Ala Pro 450 455 460 Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr His Leu 465 470 Ser Asn Gly Thr Val His Tyr Asp Val Ser Thr Lys Lys 485 490 Thr Thr Gly Thr Ile Leu Ser Ser Glu Ala Thr Ile 500 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asp Ile 515 520 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | o Gly E u Leu A s Leu T 4 e Thr T 510 e Pro V 5 | Asn Val 480 Thr Trp 495 Tyr Arg Val Asn |
| Val Lys Asn Leu Thr Gln Ser Gly Asp Val Thr His Leu 465 Ser Asn Gly Thr Val His Tyr Asp Val Ser Thr Lys Lys 485 Thr Thr Gly Thr Ile Leu Ser Ser Glu Ala Thr Ile 500 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asn Asp Ile 515 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | E Leu T 4 e Thr T 510 e Pro V 5 | 480 Thr Trp 495 Tyr Arg /al Asn |
| 15 Ser Asn Gly Thr Val His Tyr Asp Val Ser Thr Lys Lys 485 Thr Thr Gly Thr Ile Leu Ser Ser Ser Glu Ala Thr Ile 500 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asn Asp Ile 515 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | e Thr T 510 e Pro V 5 | Thr Trp 495 Tyr Arg /al Asn |
| 500 505 Ile Tyr Ala Asp Leu Asp Tyr Ile Gln Asn Asn Asp Ile 515 520 525 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | 510 e Pro V 5 p Thr A | /al Asn |
| 515 520 525 Thr Thr Ser Ala Ile Gly Pro Asp Leu Gly Gly Phe Asp | 5 p Thr A | |
| 2 111 mg mm mg mg mg mg | | |
| | | |
| Glu Ala Lys Leu Thr Tyr Thr Asn Ser Asn Gly Glu Pro 545 550 555 | | 560 |
| Leu Ile Phe Pro Arg Pro Thr Val Lys Leu Gly Tyr Gly 565 570 | 5 | 75 |
| Arg His Tyr Val Leu Val Asn Lys Asp Gl; Gln Pro Ile 580 585 | 590 | |
| Gly Thr Val Val Ser Ser Leu Ser Glu Ala His Val Leu 595 600 605 30 Asp Phe Phe Leu Pro Ser Gly Gly His Tle Val Pro | 5 | |
| 610 615 620 | | |
| Lys Leu Asp Lys Thr Thr Glu Ala Leu Gln Tyr Tyr Ser 625 630 635 Thr Asn Thr Val Ile Thr Thr Ala Asp Gly Lys Arg Tyr | | 640 |
| 35 645 650 Glu Val Pro Gly Ser Thr Pro Asn Pro Gly Gln Ile Gly | 6 | 555 |
| 660 665 Lys Lys Pro Ala Gly Asn Ala Tyr Phe Ala Tyr Lys Leu | 670 | |
| 675 680 685 40 Trp Net Gly Gly Thr Thr Asp Gln Gln Ser Glu Trp Asp | | |
| 690 695 700 Asn Trp Thr Gly Ala Gln Val Pro Leu Thr Gly Glu Asp | | |
| 705 710 715 Ala Thr Thr Glu Asn Phe Gly Ser Pro Ala Val Ala Asn | | 720 |
| 725 730 Pro Thr Thr Asn Pro Lys Ile Ile Gly Asn Leu Ile Asn | 7 | 35 |
| 740 745 Lys Asp Leu Val Val Thr Thr Ser Ser Gln Leu Thr Ile | 750 | |
| 755 760 765 50 Val Glu Asp Asn Asn Pro Asn Val Gly Thr Ile Val Val | , . Lys S | er Ser |
| 770 780 Lys Asp Asn Pro Thr Gly Thr Leu Leu Phe Ala Asn Pro | Gly T | |
| 785 790 795 Gln Asn Val Gly Gly Thr Val Glu Phe Tyr Asn Gln Gly | | |
| Ala Asp Cys Gly Met Tyr Arg Arg Ser Trp Gln Tyr Phe 820 825 | Gly I | 15 le Pro |
| 820 825 Val Asn Glu Ser Gly Phe Pro Ile Asn Asp Val Gly Gly 835 840 845 | | lu Thr |
| 60 Val Asn Gln Trp Val Glu Pro Phe Asn Gly Asp Lys Trp | Arg P | ro Ala |
| Pro Tyr Ala Pro Asp Thr Glu Leu Gln Lys Phe Lys Gly 865 870 | Tyr G | ln Ile 880 |
| Thr Asn Asp Val Gln Ala Gln Pro Thr Gly Val Tyr Ser | 8 9 | уя Gly 95 |
| Het Ile Cys Val Cys Asp Ala Phe Leu Asn Leu Thr Arg 900 905 | 910 | |
| Val Asn Tyr Ser Gly Ala Asn Leu Ile Gly Asn Ser Tyr 915 920 925 | | |
| 70 Ile Asp Ile Lys Gln Gly Ile Val Phe Pro Pro Glu Val 930 935 940 | | |
| Val Tyr Leu Phe Asn Thr Gly Thr Arg Asp Gln Trp Arg 945 950 955 | | 960 |
| Gly Ser Thr Val Ser Gly Tyr Arg Ala Gly Gln T;r Leu 965 970 | | al Pro 75 |

```
His Ser Phe Leu Val Lys Met Gln Asn Gly Ala Ser Cys Thr Leu Unk
995 1000 1005
       Ile Leu Tyr Asp Lys Leu Leu Lys Asn Thr Thr Val Asn Asn Gly Asn 1010 1015 1020
 5
       Gly Thr Gln Ile Thr Trp Arg Ser Gly Asn Ser Gly Ser Ala Asn Met 1025 1030 1035
       Pro Ser Leu Val Met Asp Val Leu Gly As<br/>n Glu Ser Ala Asp Arg Leu 1045 1050 1055
10
       Trp Ile Phe Thr Asp Gly Gly Leu Ser Phe Gly Phe Asp Asn Gly Trp 1060 1065 1070
       Asp Gly Arg Lys Leu Thr Glu Lys Gly Leu Ser Gln Leu Tyr Ala Met
1075 1080 1085
15
       Ser Asp Ile Gly Asn Asp Lys Phe Gln Val Ala Gly Val Pro Glu Leu
1090 1095 1100
       Asn Asn Leu Leu Ile Gly Phe Asp Ala Asp Lys Asp Gly Gln Tyr Thr
1105 1110
       Leu Glu Phe Ala Leu Ser Asp His Phe Ala Lys Gly Ala Val Tyr Leu
1125 1130 1135
20
       His Asp Leu Gln Ser Gly Ala Lys His Arg Ile Thr Asn Ser Thr Ser
1140 1145 1150
       Tyr Ser Fhe Asp Ala Lys Arg Gly Asp Ser Gly Ala Arg Phe Arg Leu
1155 1160 1165
25
       Ser Tyr Gly Cys Asp Glu Asn Val Asp Asp Ser His Val Val Ser Thr
1170 1175 1180
       Asn Gly Arg Glu Ile Ile Ile Leu Asn Gln Asp Ala Leu Asp Cys Thr
1185 1190 1195
       Val Thr Leu Phe Thr Ile Glu Gly Lys Leu Leu Arg Arg Leu Lys Val
1205 1210 1215
30
       Leu Ala Gly His Arg Glu Val Met Lys Val Gln Thr Gly Gly Ala Tyr
1220 1225 1230

Ile Val His Leu Gln Asn Ala Phe Thr Asn Asp Val His Lys Val Leu
               1235
       Val Glu Tyr
35
           1250
       (2) INFORMATION FOR SEQ ID NO:352
40
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 426 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
45
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
50
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...426
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352
       Thr Met Glu Val Lys Lys Asn Thr Val Val Leu Arg Leu Leu Ile Trp
                                            10
60
       Phe Val Ala IIe Leu Leu Phe His Ser Ser Arg Leu Trp Gly Gln Glu 20 25 30
       Gly Glu Gly Ser Ala Arg Tyr Arg Phe Lys Gly Phe Val Asp Thr Tyr 35 40 45
       His Ala Val Arg Ser Ser Ser Pro Phe Asp Phe Met Ser Ser Arg Thr 50 55 60
65
       Arg Val Arg Gly Glu Leu Glu Arg Ser Phe Gly Asn Ser Lys Val Ala
65 70 75 80
       Val Ser Val Asn Ala Thr Tyr Asn Ala Leu Leu Lys Asp Glu Thr Gly
85 90
70
       Leu Arg Leu Arg Glu Ala Phe Phe Glu His Gln Glu Glu His Trp Gly
100 105 110
       Leu Arg Leu Gly Arg Gln Ile Val Ile Trp Gly Ala Ala Asp Gly Val
                                   120
       Arg Ile Thr Asp Leu Ile Ser Pro Met Asp Met Thr Glu Phe Leu Ala
75
                                 135
```

```
Gln Asp Tyr Asp Asp Ile Arg Met Pro Val Asn Ala Leu Arg Phe Ser
        150 155 160

Val Phe Asn Glu Ser Met Lys Val Glu Val Val Val Leu Pro Val Phe
165 170 175

Glu Gly Tyr Arg Leu Pro Val Boo De
        Glu Gly Tyr Arg Leu Pro Val Asp Pro Arg Asn Pro Trp Asn Ile Phe
180 185 190
        Ser Leu Ser Pro Ile Ala Gln Gly Met Asn Ile Val Trp Lys Glu Glu
195 200 205
        Ala Gly Lys Pro Ala Phe Lys Val Ala Asn Ile Glu Tyr Gly Ala Arg 210 \phantom{\bigg|}215\phantom{\bigg|}220\phantom{\bigg|}
10
        Trp Ser Thr Thr Leu Ser Gly Ile Asp Phe Ala Leu Ala Ala Leu His 225 230 235 240
        Thr Trp Asn Lys Het Pro Val Ile Glu Val Gln Gly Ile Val Pro Thr
245 250 255
15
        Glu Ile Ile Val Ser Pro Arg Tyr Tyr Arg Net Gly Phe Val Gly Gly 260 265 270
        Asp Leu Ser Val Pro Val Gly Gln Phe Val Phe Arg Gly Glu Ala Ala
275 280 285
        Phe Asn Ile Asp Lys His Phe Thr Tyr Lys Ser His Ala Glu Gln Glu
290 295 300
20
        Gly Fhe Gin Thr Ile Asn Trp Leu Ala Gly Ala Asp Trp Tyr Ala Pro
305 310 315 320
        Gly Glu Trp Met Ile Ser Gly Gln Phe Ser Het Glu Ser Ile Phe Arg
325 330 335
25
        Tyr Arg Asp Phe Ile Ser Gln Arg Gln His Ser Thr Leu Ile Thr Leu 340 350
        Asn Val Ser Lys Lys Phe Phe Gly Ser Thr Leu Gln Leu Ser Asp Phe
355 360 365
        Thr Tyr Tyr Asp Leu Thr Gly Lys Gly Trp Phe Ser Arg Phe Ala Ala 370 375 380
30
       Asp Tyr Ala Leu Asn Asp Gln Ile His Leu Het Ala Gly Tyr Asp Trp 385 390 395 400

Phe Ser Ser Lys Gly Ser Gly Ile Phe Asp Arg Tyr Lys Asp Asn Ser 405 410 415
        Glu Leu Trp Phe Lys Ala Arg Tyr Ser Phe
420 425
35
        (2) INFORMATION FOR SEQ ID NO:353
40
               (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 464 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
45
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
50
                     (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...464
55
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:353
       Tyr Gly Lys Arg Arg Lys Leu Gly Thr Ser Val Arg Fro Ser Val Leu 1 5 10 15
60
       Thr Gln Ile Arg Phe Ile Leu Asp Leu His Leu Ile Thr Asp Phe Phe 20 25 30
       Glu Gly Leu Arg Val Asn Pro Ile Gly Ala Ala Ala Ile Val Ala Phe 35 \hspace{1cm} 40 \hspace{1cm} 45
       Ile Ile Asp Leu Leu Leu Cys Cys Ser Ala Phe Met Ser Ser Cys 50 55 60
65
       Glu Val Ala Tyr Phe Ser Leu Lys Pro Ile Asp Leu Gln Asn Ile Arg
65 70 75 80
       Glu Arg Asn His Ser Ser Asp Ile Ala Leu Ser Asn Leu Leu Asp Asn 85 90 95

      Ser Asn Gln
      Leu Leu Ala Thr Ile Leu Ile Gly Asn Asn Val Ile Asn 100
      105
      110

      Val Ala Ile Val Ile Leu Ser Asn Tyr Ala Ile Glu Gln Thr Phe Val 115
      120
      125

70
        Phe Ser Ser Pro Ile Ile Gly Phe Leu Ile Gln Thr Ile Leu Leu Thr
75
```

```
Thr Val Leu Leu Eu Phe Gly Glu Ile Leu Pro Lys Val Tyr Ala Arg 145 150 155 160
       Lys Asn Pro Leu Gln Tyr Ser Arg Phe Ser Ala Ala Ala Met Ser Val
165 170 175
 5
       The Tyr Lys lie Leu Ser Pro Phe Ser Lys Leu Leu Val Lys Ser Thr 180 185 190
       Gly Ile Val Thr Arg Gly Ile Ser Lys Lys Lys Tyr Asp Met Ser Val
195 200 205
       Asp Glu Leu Ser Lys Ala Val Ala Leu Thr Thr Thr Glu Gly Glu Pro
210 215 220
10
       Glu Glu Lys Glu Met Ile Asn Glu Ile Ile Lys Phe Tyr Asn Lys Thr
225 230 235 240
       Ala Cys Glu Ile Met Val Pro Arg Ile Asp Ile Val Asp Val Asp Leu
245 250 255
15
       Ser Trp Pro Phe Arg Lys Met Leu Asp Phe Val Val Ser Ser Gly Tyr
260 265 270
       Ser Arg Leu Pro Val Ser Glu Gly Ser Glu Asp Asn Ile Lys Gly Val
275 280 285
       Ile Tyr Ile Lys Asp Leu Ile Pro His Met Asp Lys Gly Asp Glu Phe
290 295 300
20
       Asp Trp His Pro Leu Ile Arg Lys Ala Tyr Phe Val Pro Glu Asn Lys 305 310 315
       Arg Ile Asp Asp Leu Leu Glu Glu Phe Arg Ala Asn Lys Val His Val
325 330 335
25
       Ser Ile Val Val Asp Glu Phe Gly Gly Thr Cys Gly Leu Ile Thr Het 340 345
       Glu Asp Ile Leu Glu Glu Ile Val Gly Glu Ile Thr Asp Glu Tyr Asp
355 360 365
       Glu Glu Glu Leu Pro Phe Lys Val Leu Gly Asp Gly Ser Tyr Leu Phe
370 375 380
30
       Glu Gly Lys Thr Ser Leu Ser Asp Val Arg His Tyr Leu Asp Leu Pro 385 390 395 400
Glu Asn Ala Phe Gly Glu Leu Gly Asp Glu Val Asp Thr Leu Ser Gly 405 410 415
       Leu Phe Leu Glu Ile Lys Gln Glu Leu Pro His Val Gly Asp Thr Ala
420 425 430
35
       Val Tyr Glu Pro Phe Arg Phe Gln Val Thr Gln Met Asp Lys Arg Arg
435 440 445
        Ile Ile Glu Ile Lys Ile Phe Pro Phe Glu Arg Thr Trp Glu Val Glu
40
        (2) INFORMATION FOR SEQ ID NO:354
              (i) SEQUENCE CHARACTERISTICS:
45
                   (A) LENGTH: 266 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
50
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
55
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature (B) LOCATION 1...266
60
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:354
       The Ile Tyr Arg Ser Thr Met Lys Leu Leu Leu Tyr Leu Leu Val
                         5
       Leu Ser Thr Leu Ser Pro Met Tyr Ser Gln Het Leu Phe Ser Glu Asn 20 25 30
65
       Leu Thr Het Asn Ile Asp Ser Thr Lys Thr Ile Gln Gly Thr Ile Leu 35 40
       Pro Val Leu Asp Phe Lys Thr Glu Lys Glu Asn Val Phe Thr Phe Lys 50 55 60
70
       Asn Thr Ala Asn Leu Asn Leu Leu Ile Lys His Gly Gln Val Ile Asn 65 70 80
       Leu Ile Asn Lys Leu Glu Phe Ser Thr Tyr Gly Asn Lys Val Thr Val
                        85
                                              90
       Ser Gly Gly Tyr Val His Thr Glu Tyr Arg Tyr Leu Leu His His Val
75
                                            105
```

```
Phe Glu Val Tyr Pro Tyr Val Glu Ser Gln Trp Ala Glu Ser Arg Gly
              115
                                      120
       Het Lys Tyr Lys Val Ser Thr Gly Leu Gln Ser Arg Tyr Arg Leu Val
                            135
                                                       140
 5
       Asn Ser Asp Asn Cys Leu Met Phe Ala Thr Leu Gly Val Phe Phe Glu
145 150 155 160
       Phe Glu Lys Trp Glu Gln Pro Ala Thr Ser Leu Phe Ala Gly Thr Tyr
165 170 175
       Ala Tyr Ser Arg Ser Ile Lys Ser His Leu Ser Ile Ser Phe Arg His 180 $180$
10
       Arg Leu Gly Glu His Trp Glu Phe Thr Thr Thr Ala Ile His Gln Gly
195 200 205
       Lys Pro Asp Ser Tyr Phe Lys Lys Ala Arg Phe Gly Gly Ala Ile Asp 210 215 220
15
       Leu Lys Tyr His Ile Thr Pro Thr Ile Gly Ile Arg Gly Ala Tyr Arg 225 230 235 240
       Ile Ile Tyr Asp Thr Ala Pro Ile Val Pro Val Arg Lys Asp Tyr Asn
245 250 255
       Thr Val Asp Val Gly Ile Asp Ile Ser Phe
20
       (2) INFORMATION FOR SEQ ID NO: 355
            (i) SEQUENCE CHARACTERISTICS:
25
                  (A) LENGTH: 907 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
30
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
35
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...907
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:355
      Thr Ile Cys Val Ala Phe Leu Ser Ala Pro Val Ala Ala Leu Arg Glu 1 5 10 15
      Ser Pro Pro Met Gly Ala Glu Arg Lys Thr Pro Ser Leu Leu Pro Leu
20 25 30
45
      Leu Phe Glu His Ser Asp Lys Gly Lys Gly Phe Asp Tyr Arg Leu Phe 35 40 45
      Thr Ser Asn Lys Leu Lys Val Phe Ala Thr Gly Asn Ser Arg Tyr Ile
50 60
50
      His Asn Lys Pro Thr Ile Gln Ala Met Lys Arg Ile Val Leu Ser
65 70 80
      Ser Phe Leu Phe Val Leu Ser Ile Leu Ser Leu Net Ala Gln Asn Asn 85 90 95
      Thr Leu Asp Val His Ile Ser Gly Thr Ile Lys Asp Ala Ser Ser Gly 100 105 110
55
      Glu Pro Val Pro Tyr Ala Thr Val Ser Ile Arg Leu Thr Gly Ala Asp
115 120 125
      Thr Thr Gln Val Phe Arg Gln Val Thr Asp Gly Asn Gly Tyr Phe Val
130 135 140
60
      Ile Gly Leu Pro Ala Ala Pro Ser Tyr His Leu Thr Ala Ser Phe Val
145 150 155 160
      Gly Het Lys Thr His Thr Het Gln Ile Ser Arg Gly Asn Gly Gln His
165 170
      Asp Ile Lys Ser Ile Asp Ile Ser Leu Glu Ser Glu Asp Lys Gln Leu
180 185 190
65
      Ser Thr Val Thr Val Ser Ala Ala Arg Pro Leu Val Lys Met Glu Ile
195 200 205
      Asp Arg Leu Ser Tyr Asn Met Lys Asp Asp Pro Ala Ala Lys Thr Asn 210 215 220
70
      Asn Leu Leu Glu Met Leu Arg Asn Val Pro Leu Val Thr Val Asp Gly 225 230 235 240
      Gln Gl; Asn Ile Gln Val Lys Gly Ser Ser Asn Phe Lys Ile His Leu
245 250 255
                                           250 255
      Asn Gly Arg Pro Ser Thr Met Val Ser Ser Asn Pro Lys Glu Val Phe
75
```

| | Arg | Ser | 11∈ 275 | Pro | Ala | His | Thr | Ile 280 | Lys | Arg | Val | Glu | Val 285 | Ile | Thr | Asp |
|-----------|------------|------------|------------|-------------|------------|-----------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Pro | Gly 290 | | Lγε | Tyr | Asp | Ala 295 | | Gly | Thr | Ser | Ala 300 | | Leu | Asp | Ile |
| 5 | Val 305 | Thr | Glu | Glu | Gly | Lys 310 | | Leu | Glu | Gly | Туг 315 | - | Gly | Ser | Ile | Thr 320 |
| | Ala | Ser | Val | Ser | Asn 325 | Asn | Pro | Thr | Ala | Asn 330 | Gly | ser | Ile | Phe | Leu 335 | Thr |
| 10 | | | | 340 | | | | | Thr 345 | | | | | 350 | _ | - |
| | | | 355 | | | | | 360 | Phe | | | | 365 | | | |
| 15 | | 370 | | | | | 375 | | Gly | | | 380 | | | | |
| 15 | 385 | | | | | 390 | | | Glu | | 395 | | | | | 400 |
| | | | | | 405 | | | | Trp | 410 | | | | | 415 | |
| 20 | | | | 420 | | | | | Ser 425 | | | | | 430 | | - |
| | | | 435 | | | | | 440 | Ala | _ | | | 445 | | | |
| 25 | | 450 | | | | | 455 | | Asn | _ | | 460 | | | | |
| 20 | 465 | | | | | 47û | | | Ala | | 475 | | | | • | 480 |
| | | | | | 485 | | | | Met | 490 | | | | | 495 | - |
| 30 | | | | 50 0 | | | | | 505 Ala | - | | | | 510 | | |
| | | | 515 | | | | | 520 | Ser | | | | 525 | | _ | |
| 35 | | 530 | | | | | 535 | | Pro | | | 540 | | | | |
| | 545 | | | | | 550 | | | Asp | | 555 | | | | | 560 |
| | | | | | 565 | | _ | | Gln | 570 | _ | | - | | 575 | - |
| 40 | | | | 580 | | | | | 585 Ala | - | | | | 590 | _ | |
| | | | 595 | | | | | 600 | Asp | | | | 605 | | | |
| 45 | | 610 | | | | | 615 | | Gln | | | 620 | | | | |
| | 625 Arg | Ile | Gln | Arg | | 630 Ala | Ile | Gly | Gln | | 635 Asn | Pro | Tyr | Arg | Leu | 640 Gln |
| F.O. | Thr | Asn | Asp | | 645 Gln | Val. | Gln | Tyr | Gly | 650 Asn | Pro | Агр | Leu | Lys | 655 Ser | Glu |
| 50 | Lys | Arg | | 660 His | Val | G1 _Y | Leu | | 665 Tyr | Asn | Gln | Туг | | 670 Ala | Lys | Val |
| | Met | Leu | 675 Thr | Ala | Ser | Leu | Asp | 680 Tyr | Asp | Phe | Суз | | 685 Asn | Ala | Ile | Gln |
| 55 | Asn 705 | 690 Tyr | Thr | Phe | | | 695 Pro | Ala | Asn | Pro | | 700 Leu | Phe | His | Gln | |
| | | Gly | Asn | Ile | | 710 Arg | Glu | His | Ser | Phe 730 | 715 Ser | Leu | Asn | Thr | | 720 Ala |
| 60 | Het | Tyr | Thr | Pro 740 | | Val | Trp | Val | Arg 745 | | Net | Leu | Asn | Gly 750 | 735 Asn | Ile |
| 00 | Asp | Arg | Thr 755 | | Gln | Lys | Ser | Glu 760 | Ala | Leu | Glγ | Ile | Asp 765 | | Asn | Ser |
| | Trp | Ser | | Het | Val | Tyr | Ser 775 | | Leu | Het | Phe | Thr 780 | | Pro | Lys | Asp |
| 65 | Trp 785 | | Val | Asn | Leu | Phe 790 | | Gly | Tyr | Tyr | His 795 | | Gly | Arg | Ser | Tyr 800 |
| | | Thr | Lys | туг | Asp 805 | | Asn | Val | Phe | Asn 810 | | Ile | GJ 7. | Ile | Ala 815 | |
| 70 | Gln | Leu | Phe | Asp 820 | | Lys | L≘u | Arg | Val 825 | | Leu | Ser | Ala | Aen 830 | | Ile |
| | His | Ala | Lys 835 | | Ser | Thr | Trp | Lys 840 | Ser | Arg | Thr | Ile | Gly 845 | | Gly | Phe |
| | Thr | Ile 850 | | Ser | Glu | Asn | Ala 855 | | Ile | Gln | Arg | Ser 860 | | Ser | Leu | Ser |
| 75 | Leu | | Tyr | Ser | Phe | G17. | | Het | Asn | Thr | Gln | | Arg | Lys | Val | Glu |

```
870
                                                    875
                                                                         880
       Arg Thr Ile Val Asn Asp Asp Leu Lys Gln Thr Ser Ser Gln Gly Gln
885 890 895
                                            890
       Gln Gly Gly Gln Gly Asn Pro Thr Gly Asn
  5
                    900
       (2) INFORMATION FOR SEQ ID NO: 356
             (i) SEQUENCE CHARACTERISTICS:
10
                  (A) LENGTH: 450 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
15
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
20
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...450
25
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:356
       Ala Gln Gly His Gln Gln Gly Arg Ser Asp Gln Asp Arg Leu Pro Pro
25 30
30
       Val Arg Gly Tyr Arg Tyr Ala Arg Cys Leu Gly Arg Tyr Glu Gly Phe
35 40
       Gly Glu Arg Net Arg Leu Ser Ala Ile Leu Ile Ala Leu Ile Val Net
50 55 60
35
       Leu Pro Ala Val Leu Ser Gly Gln His Tyr Tyr Ser Met Ala Gly Glu
65 70 75
       Arg Leu Glu Thr Asp Ser Ile Arg Pro Asn Glu Leu Ser Ala Ser Ile
85 90 95
       Arg Ser Ala Leu Phe Phe Arg Asn Asn Glu Tyr Asn Ala Arg Ser Val
40
       Lys Gly Tyr Thr Leu Pro Gly Ala Arg Val Ser Ala Phe Ala Ser Tyr 115 120 125
       Ser Leu Pro Ala Ala His Gly Val Lys Leu Ser Leu Gly Val Ser Thr
130 135 140
45
       Leu Asn Tyr Trp Gly Ala Ser Arg Tyr Pro Ala Gly Ile Ala Tyr Ser
145 150 155 160
       Asp Leu Pro Tyr Trp Thr Asp Tyr Asn Asp Tyr Val Arg Leu Arg Ile
165 170
       Leu Pro Tyr Val Gln Ala Met Leu Lys Pro Thr Ala Thr Thr Ala Leu
180 185 190
50
      Het Leu Gly Asn Ile Ala Gly Gly Thr Ala His Gly Leu Ile Glu Pro
195 200 205
      Ile Tyr Asn Pro Glu Leu Asp Leu Thr Ala Asp Pro Glu Ala Gly Val
210 215 220
      Gln Phe Arg Gly Asp Trp Thr Arg Phe Arg Het Asp Val Trp Val Asn 225 230 235 240
55
      Trp Met Ser Met Ile Phe Lys Asn Asp Asn His Gln Glu Ser Phe Val 245 250 255
       Phe Gly Leu Ser Thr Thr Ser Lys Leu Leu Ser Gly Glu Gly Lys Trp
260 265 270
60
      Arg Leu Glu Leu Pro Leu Gln Ala Ile Ala Thr His Arg Gly Gly Glu 275 280 285
      Tyr Asn Trp Ala Gln Gln Asp Thr Val His Thr Trp Val Asn Gly Ala
290 295 300
      Val Gly Leu Lys Leu Ser Tyr Arg Pro Arg Thr Asp Lys Pro Met Gln 305 310 315 320
65
      The Trp Gly Ser Ala Tyr Gly Val Ala Ala Leu Ser Ser Gly Gly Tyr 325 330 335
      Phe Pro Tyr Glu Arg Gly Trp Gly Gly Tyr Leu Ser Leu Gly Met Asp 340 345
70
      Leu Glu His Phe Ala Phe Arg Thr Asp Tyr Trp Tyr Gly Arg His Tyr 355 360 365
      Val Ser Pro Phe Ala Ala Pro Phe Ala Asn Ser Leu Thr Tyr Asp Lys
370 375 380
75
      Gin Pro Leu Thr Asn Gly Trp Gly Asp Tyr Ile Arg Leu Tyr Ala Asp
```

```
390
       Tyr Ser Trp Arg Met Ala Arg Ser Val Ser Leu Ala Ala Val Ala Arg
405 410 415
       Val Trp Phe Gln Pro Ser Asp Arg Phe Ala Het Ser His Ala Leu Glu
420 425 430
 5
        Leu Thr Met Arg Ile Asp Pro Lys Phe Pro Ile Ala Phe Leu Lys Gly
10
        (2) INFORMATION FOR SEQ ID NO:357
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 447 amino acids
15
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
20
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
25
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...4\overline{47}
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:357
30
       Pro Phe Ser Pro Pro Trp Val Cys Ala Gly Ala Ala Ser Ser Val Arg
                                                 10
       Thr Arg Pro Cys Ser Thr Val Ser Ser Thr Ala Ala Ser Ser His Arg
20 25 30
35
       Arg Ser Glu His Met Arg Ser Ala Asp Ser Ser Pro Ala Tyr Ser Pro 35 40 45
       Ile Gly Ser Leu Ser Ser Ser His Ser Phe Asp Ser Ile Pro Asp Gln 50 55 60
       Arg Het Asn Lys Ser Leu Leu Ser Leu Ala Cys Leu Ile Leu Cys Gly 65 70 75 80
40
       Met Pro Ala Ile Ala Gln Gln Thr Gly Pro Ala Glu Arg Ser Gly Glu
85 90 95
       Pro Ser Leu Ala Glu Arg Val Phe Gly Leu Glu Gln Lys Gln Lys Lys
100 105 110
45
       Leu Lys Val Tyr Leu Gly Ile Gln Ser Phe Tyr Asp Gln Pro Leu Val 115 120 125
       Asp Asp Glu Ser His Ile Gly His Phe Lys Val Gln Glu Leu Arg Het
130 135 140
       Ser Ala His Gly Glu Leu Asn Arg His Leu Ser Phe Asp Trp Arg Gln 145 150 155 160
50
       Arg Leu Asn Arg Ala Ala Asp Gly Thr Ser Phe Ala Asp Asn Leu Ser
165 170 175
       Asn Ala Ile Asp Ile Ala Gly Val Asp Trp His Pro Asn Asp Lys Val
       Ser Phe Phe Phe Gly Arg Gln Tyr Ala Arg Phe Gly Gly Ile Glu Tyr
195 200 205
55
       Asp Het Asn Pro Val Glu Ile Tyr Gln Tyr Ser Asp Leu Val Asp Tyr 210 215 220
       Met Thr Cys Tyr Thr Ser Gly Val Asn Phe Ala Trp Asn Phe His Pro 225 230 235 240
60
       Glu Gln Gln Leu Gln Leu Gln Val Leu Asn Ala Tyr Asn Asn Arg Phe
245 250 255
       Ala Asp Arg Tyr His Val Thr Pro Asp Val Ala Thr Ala Thr Ser Tyr
260 265 270
65
       Pro Leu Leu Tyr Ser Ala Gln Trp Asn Gly Thr Leu Leu Gly Gly Ala
275 280 285
       Leu His Met Arg Tyr Ala Val Ser Met Ala His Gln Ala Gln Glu Arg
290 295 300
       Asn Met Trp Tyr Phe Thr Ala Gly Asn Leu Phe Asn Pro Gly Lys Arg
305 310 315 320
Ile Asn Gly Tyr Leu Asp Leu Thr Tyr Ser Ile Glu Gly Leu Asp Asp
325 330 335
70
       Lys Gly Ile Met Thr Ala Arg Tyr Gly Lys Gly Lys Thr Leu Thr Asp
                     340
                                           345
75
       Val Lys Tyr Tyr Ala Leu Val Ser Lys Trp Asn Phe Arg Ile Phe Asp
```

```
355
                                       360
        Gln Val Asn Leu Phe Leu Lys Gly Met Tyr Glu Asn Gly Tyr Ala Pro
370 375
        Ala Gln Tyr Gly Glu Ser Ser His Thr Arg His Ser Tyr Gly Tyr Met
385 390 395 400
  5
        Gly Gly Val Glu Tyr Tyr Pro Thr Glu Thr Asn Phe Arg Leu Phe Val
405 415
        Thr Tyr Ile Gly Arg His Tyr Arg Tyr Ser Ala Thr Glu Thr Glu Ser
420 425 430
 10
        Thr Asn Ala Leu Arg Ala Gly Leu Ile Tyr Gln Ile Pro Phe Leu
        (2) INFORMATION FOR SEQ ID NO:358
 15
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 227 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 20
            (ii) HOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
25
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...227
30
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:358
       Tyr Lys Gin Ala Ile Met Lys His Leu Phe Lys Ser Thr Leu Val Leu
                                               10
                                                                    15
35
       Leu Cys Ala Leu Ser Phe Ser Gly Thr Tyr Thr Phe Ala Gln Glu Asn
                                          25
       Asn Thr Glu Lys Ser Arg Phe Asp Phe Ser Val Arg Leu Gly Gln Gly
35 40 45
       Tyr Ile Ala Gly Ser Thr Thr Asn Leu Met Tyr Gly Tyr Thr Ser Ala 50 60
40
       Asn Asp Arg Leu Leu Ser Gly Ala Ile Tyr Leu Gly Leu Thr Pro Ser 65 70 75
       Lys Lys Glu Asn Ala Thr Gly Val Ala Phe Arg Phe Leu Ser Pro Ser
85 90 95
       85 90 95
Pro Gly Tyr Tyr Val Asp Ile Ser Gly Lys Glu Asn Thr Leu Asn Tyr
100 105 110
45
       Ala Phe Tyr Val Val Gly Ala Tyr Asn Arg Ile Ala Ile Pro Ile Arg
115 120 125
       Pro Ile Lys Asn Phe Asn Phe Ile Phe Ser Thr Glu Val Gly Met Ala
130 135
50
       Trp Met Ser Arg His Glu Gln Ile Tyr Asn Ser Thr Ser Gln Thr Trp 145 150 155 160
       Asp Lys Gln Arg Lys Ser Arg Ser Gly Leu Asp Phe Gly Leu Gly Met 165 170 175
       His Leu Gln Unk His Ile Asn Lys Thr Val Tyr Phe Met Ala Gly Thr
180 185 190
55
       Asp Leu Thr Ser Cys Met Phe Gly Lys Arg Ile Asn Asp Tyr Gln Gln 195 200 205
       Lys Asp Arg Thr Phe Ile Ala Leu Ile Asp Asn Ser Ile Gly Ile Gly
60
       Leu Asn Leu
       225
       (2) INFORMATION FOR SEQ ID NO:359
65
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 406 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
70
           (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
75
           (vi) ORIGINAL SOURCE:
```

```
(A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
 5
                  (B) LOCATION 1...406
           (::i) SEQUENCE DESCRIPTION: SEQ ID NO:359
      10
      Ile Ile Lys Lys Met Leu Lys Asn Lys Leu Ala Pro Leu Ala Ile Leu
20 25 30
       Phe Leu Phe Ala Pro Lys Ala Met Lys Ala Gln Glu Gln Leu Asn Val
35 40
15
      Val His Thr Ser Val Pro Ser Leu Asn Ile Ser Pro Asp Ala Arg Ala
50 55 60
      Ala Gly Met Gly Asp Ile Gly Val Ala Thr Thr Pro Asp Ala Tyr Ser 65 70 75 80
      Gln Tyr Trp Asn Pro Ser Lys Tyr Ala Phe Met Asp Thr Lys Ala Gly
85 90 95
20
      Ile Ser Phe Ser Tyr Thr Pro Trp Leu Ser Lys Leu Val Asn Asp Ile
100 105 110
      Ala Leu Met Gln Met Thr Gly Phe Tyr Lys Leu Gly Thr Asp Glu Asn
115 120 125
25
      Gln Ala Ile Ser Ala Ser Leu Arg Tyr Phe Thr Leu Gly Lys Leu Glu 130 135 140
      Thr Phe Asp Glu Leu Gly Glu Ser Met Gly Glu Ala His Pro Asn Glu
145 150 155 160
      Phe Ala Val Asp Leu Gly Tyr Ser Arg Gln Leu Ser Glu Asn Phe Ser
165 170 175
30
      Net Ala Val Ala Leu Arg Tyr lle Arg Ser Asp Gln Ser Thr His Asn
180 185 190
      Thr Gly Glu Asn Gln Ala Gly Asn Ala Phe Ala Ala Asp Ile Ala Gly
195 200 205
35
      Tyr Leu Gln Lys Tyr Val Leu Leu Gly Asn Ala Glu Ser Leu Trp Ser
210 215 220
      Leu Gly Phe Asn Val Lys Asn Ile Gly Thr Lys Ile Ser Tyr Asp Gly 225 230 230 235
      Gly Val Thr Ser Phe Phe Ile Pro Thr Ser Leu Asn Leu Gly Thr Gly
245 250 255
40
      Leu Leu Tyr Pro Ile Asp Asp Tyr Asn Ser Ile Asn Phe Asn Leu Glu
260 265 270
      Leu Ser Lys Leu Leu Val Pro Thr Pro Pro Ile Met Asp Gln Asn Asp 275 280 285
45
      Gln Ala Gly Tyr Glu Ala Ala Leu Lys Lys Tyr Gln Glu Thr Ser Ser
290 295 300
      Ile Ser Gly Ile Phe Ser Ser Phe Gly Asp Ala Pro Gly Gly Leu Lys 305 310 315 320
      Glu Glu Phe Arg Glu Ile Thr Trp Gly Leu Gly Ala Glu Tyr Ser Tyr 325 330 335
50
      Asp Asp Lys Phe Phe Val Arg Ala Gly Tyr Ser Tyr Leu His Pro Thr 340 345 350
      Lys Gly Asn Leu Gln Tyr Phe Thr Ala Gly Ala Gly Phe Lys Met Asn 355 360 365
55
      Ile Phe Arg Ile Asp Ala Ser Tyr Leu Leu Ser Thr Ile Gln Ser Asn 370 375 380
       Pro Leu Asp Gln Thr Leu Arg Phe Thr Leu Ala Phe Asp Met Asp Gly
      Leu Arg Asn Leu Phe His
60
       (2) INFORMATION FOR SEQ ID NO: 360
            (i) SEQUENCE CHARACTERISTICS:
65
                 (A) LENGTH: 452 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
70
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
75
```

```
(ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                      (B) LOCATION 1...452
  5
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360
        Leu Ile Ile Glu Lys Glu Met Lys Thr Thr Val Gln Gln Ile Ile Leu 1 5 . 10 15
        Cys Leu Ala Leu Met Met Ser Gly Val Leu Gly Gly Asn Ala Gln Ser 20 25 30
 10
         Phe Trp Glu Glu Ile Ala Pro Pro Phe Ile Ser Asn Glu Pro Asn Val
35 40
         Lys Tyr Ile Ile Pro Asn Met Gly Ile Asp Ser Lys Gly Thr Ile Tyr 50 60
15
        Val Thr Val Thr Lys Arg Ile Gln Gln Gly Ala Asn Tyr Thr Ser Glu 65 70 75
        Gln Leu Gly Het Tyr Tyr Arg Pro Leu Gly Asp Asn Glu Gln Trp Trp
85 90 95
        Lys His Asp Pro Tyr Phe Asp Asp Lys Ile Val Ala Asp Ile Gln Thr 100 105 110
20
        Asp Ala Tyr Gly Arg Val Tyr Val Cys Thr Thr Ser Ser Arg Asp Gln 115 120
        Glu Tyr Gln Leu Tyr Ile Asn Glu Gln Asn Glu Trp Arg Cys Ile Phe
130 135 140
        Lys Thr Ser Val Ser Thr Tyr Glu His Gly Het Ala Val Phe Arg Ser
145 150 150
25
        Ser Thr Gly Val Thr Tyr Ile Gly Thr Arg His His Ile Phe Ala Ser
165 170
        Gly Val Asn Asp Phe Glu Phe Asn Thr Ile Tyr Glu Asp Ser Thr Pro
180 185 190
30
        Met Ser Cys Arg Phe Ala Glu Ala Thr Asn Ser Gly Thr Ile Tyr Leu
195 200 205
        Ala Leu Met His Glu Thr Thr Het Ser Thr Thr Ile Leu Thr Tyr Gln
210 215 220
        Asn Gly Glu Phe Val Asp Ile Ser Glu Ser Glu Leu Ser Asn Ser Ile 225 230 235 240 Ile Ala Ser Het Cys Ser Asn Lys Glu Gly Asp Ile Ile Ala Leu Val 245 250 250 255
35
        Thr Ser Tyr Thr Gly Phe Met Ser Gly Thr Leu Ala Ile Arg Lys Ala 260 265 270
40
        Asp Glu Gly Lys Trp Gln Leu Val Gly Gly Asp Ile Gln Asn Ala Ile
275 280 285
        Val Gln Asn Ile Cys Met Met Asp Asp Asn Lys Ile Ala Cys Glu Val 290 295 300

Phe Gly Thr Pro Asn Gly Val Asp Gly Arg Thr Arg Val Cys Val Ser 310 310 315 320
45
        Asp Ala Ser Val Phe Asp Phe Glu Trp Tyr Glu Asp Glu Ile Tyr Gly
325 330 335
        Gly Leu Ile Phe Asp Thr Phe Phe Tyr Ser Pro Trp Asp Lys Leu Leu 340 345 350
50
        Tyr Ala Lys Phe Gly Gly Ile Met Leu Arg Ser Lys Glu Ser Phe Ile
355 360 365
        Thr Ser Phe Ile Ser Pro Thr Val Val Gln Gly Val Asp Val Tyr Thr 370 375 380
55
        Leu Ala Gly Lys Ile Arg Ile Glu Ser Glu Thr Pro Val Ser Glu Val 385 390 395 400
       Leu Leu Phe Asp Leu Ala Gly Arg Met Val Leu Arg Gln Thr Ile Asp 405 415

Asn Lys Ile Tyr Ser Asp Ile Asp Thr Asn Gly Leu Lys Arg Ser Gly 420 425 430
60
        Ile Tyr Val Val Ser Val Arg Leu Ser Ser Gly Gln Val Phe Ser His
        Lys Val Gln Val
65
        (2) IMFORMATION FOR SEQ ID NO: 361
```

- (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 331 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- 75 (iii) HYPOTHETICAL: YES

70



| | <pre>(vi) ORIGINAL SOURCE: (A) ORGANISM: Porphyromonas gingivalis</pre> | | | | | | | | | | | | | | | |
|------------|--|----------------|------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|----------------|
| 5 | | (ix | | A) N | | | | | ature | 3 | | | | | | |
| 10 | | (xi |) SE(| QUENC | CE DI | ESCR | PTI | ON: 8 | SEQ I | ED NO | D:36 | 1 | | | | |
| | Gly 1 | Leu | T).r | Gly | Met 5 | Ser | Val | Val | Pro | 11e 10 | Ile | Ile | Tyr | Leu | Cys | Gl |
| | | Ser | Asn | Tyr 20 | Ala | Arg | Leu | Met | 11e 25 | | Arg | Сув | Leu | Ile 30 | | Arg |
| 15 | Pro | Arg | Thr 35 | | Leu | Phe | Gly | Leu 40 | | Phe | Val | Val | Gly 45 | | Phe | Ser |
| | Ala | Het 50 | | Gln | Glu | Lys | Lys 55 | | Ser | Leu | Ser | Thr 60 | | Gln | Pro | Val |
| 20 | Pro | Asn | Ser | Ser | Met | Val 70 | | Gln | Thr | Pro | Leu 75 | | Ser | Ile | Asp | His 80 |
| | Pro | Val | Leu | Pro | Ala 85 | | Phe | Gln | Asn | Thr 90 | | Thr | Leu | Lys | Arg 95 | |
| | Arg | Asp | rys | His 100 | | Ser | Asp | Ala | Leu 105 | | Asn | Gly | Leu | Lys 110 | | His |
| 25 | Arg | Ser | Ser 115 | | Gln | Leu | Asn | Glu 120 | | Leu | Asn | Phe | Ala 125 | | Glu | Arg |
| | Arg | Asp 130 | | Val | Ser | Pro | Leu 135 | | Gln | Thr | Arg | His 140 | | Ala | Gly | Val |
| 30 | Leu 145 | Ser | Trp | Arg | Pro | Thr 150 | | Arg | Met | His | Phe 155 | | Thr | Ser | Gly | Asn 160 |
| | | G17. | Leu | Gly | His 165 | | Leu | Leu | Thr | Gly 170 | | Arg | Lys | Asp | Phe 175 | |
| | Trp | Asn | Ala | Gly 180 | _ | Asp | Phe | Leu | Leu 185 | | Gln | Asn | Leu | Thr 190 | | His |
| 35 | Val | Gln | Gl _Y 195 | Gly | Trp | Gln | Gln | Asn 200 | | Gly | Phe | Ile | Pro 205 | | Thr | Ala |
| | Val | Asn 210 | Gly | Gln | Leu | Arg | Trp 215 | Gln | Ala | Thr | Glu | Arg 220 | Leu | Ser | Phe | Thr |
| 4 0 | Thr 225 | Glγ | Ile | Asp | Tyr | Arg 230 | Gln | Val | Gln | Trp | Asn 235 | | Phe | Asp | Asn | Arg 240 |
| | Thr | Phe | Ser | Leu | Lys 245 | Gly | Ser | Ala | Arg | Tyr 250 | Glu | Val | Met | Asp | Asn 255 | Vāl |
| | Phe | Val | Asn | Gly 260 | Phe | Gly | Ser | Tyr | Pro 265 | Leu | Tyr | Ser | Ser | Thr 270 | Arg | Ser |
| 45 | Gly | Leu | Asn 275 | Met | Ala | Vāl | Pro | Het 280 | His | Glу | Phe | Gly | Pro 285 | Gln | Туr | Gly |
| | Gly | Ser 290 | Leu | Glu | Leu | Lys | Val 295 | Ser | Glu | Arg | Phe | Gly 300 | Phe | Ala | Val | Gly |
| 50 | Met 305 | Glu | Arg | Glu | Tyr | Asn 310 | Ile | Тгр | Thr | Arg | Arg 315 | Trp | Glu | Thr | His | Tyr 320 |
| | Phe | Ala | туг | Pro | Val 325 | Phe | Tyr | Gly | Asp | Lys 330 | Lys | | | | | |
| | (2) | INFO | RI:IA | NOI | FOR | SEQ | 1 DI | 10:36 | 52 | | | | | | | |
| 55 | | (i) | SEC | | | | | | | | | | | | | |
| | | | (E | 3) Ti | PE: | amir | no ac | cid | acio | is | | | | | | |
| 60 | | ,,,, | | | POLC | | | | | | | | | | | |
| | | | HOI | | | | - | ein | | | | | | | | |
| 65 | , | (111) (111) | ORI | | | | | | | | | | | | | |
| 00 | | (\ 1) | | | | | | hyro | mona | s gi | ngiv | alis | 3 | | | |
| | | (ix) | FEA | | | ev. | mico | . f | ture | | | | | | | |
| 70 | | | | | CATI | | | | icur e | ; | | | | | | |
| | | (xi) | SEÇ | MENC | E DE | SCRI | PTIC | M: 5 | EQ I | D NO | 362 | 2 | | | | |
| 75 | Glu 1 | Thr | Asn | Ser | Trp 5 | Val | Ser | Ser | Asp | Cys 10 | Asn | Ser | Thr | Thr | Иеt 15 | Lys |

```
Thr Asn Arg Arg Tyr Ala Phe Val Leu Pro Leu Leu Leu Leu Thr Gly 20 25 30
        Leu Leu Ala Trp Gly Gln Asp Ser Ser His Gly Ser Asn Thr Ala Phe
35 40 45
  5
        Ala Thr Asp Ser Ser Ser Arg Glu Leu Pro Thr Glu Gln Ser Ala Tyr
50 55 60
        Arg Ile His Ser Ala Tyr Met Val Gly Gly Gly Gly Ser Ile Thr Arg 65 70 75 80
        Asp Thr Tyr Leu Ser Pro Leu Arg Tyr Gly Gly Trp Thr Leu Asn Leu
85 90
 10
        Leu Gly Glu Lys Thr Phe Pro Leu Lys Ala Ser Asp Ser Arg Trp Met
100 110

      Ile Arg Thr Gly His Glu Leu Asp Phe Ala Leu Met Asp Asn Fro Ala 115 .
      120 .
      125 .

      Asn Asn Ala His Phe Tyr Ser Leu Leu Tyr Asn Gly Ser Ala Ala Ala 130 .
      135 .
      140 .

 15
        Leu Tyr Arg Leu Gly Ala Lys His Leu Arg Ala Ala Trp Met Asp Asn
145 150 155 160
        Leu Arg Leu Ala Phe Gly Pro Gly Leu Glu Ile Gly Leu Gly Gly Ile
165 170 175
20
        Tyr Ser Thr Arg Asn Gly Asn Asn Pro Ala Thr Leu Lys Leu Tyr Thr
180 185 190
        Asn Ala Ile Ala Gln Ala Ser Ile Gly Tyr Tyr Val Pro Ser Glu Thr 195 200 205
        Phe Pro Leu Tyr Phe Arg Leu Leu Ser Gln Ile Asn Leu Phe Gly Ile
210 215 220
25
        Ala Tyr Gly Asn Gly Phe Gly Glu Ser Tyr Tyr Glu Asn Phe Leu Leu
225 230 235 240
       Asn Asn Gly Ile Ala Gly Ser Leu His Phe Thr Tyr Pro Gly Lys Phe 245 250 255
30
       Thr Arg Phe Thr Thr Leu Ile Thr Ala Asp Ile Pro Ile Arg Asn Phe 260 265 270
       Cys Thr Leu Arg Val Gly Tyr Arg Tyr Ser His Leu Gly Ser Ser Leu 275 280 285
       Asn Ala Leu Asp Thr Arg Ile His Ser His Thr Ala Phe Ile Gly Phe 290 295 300
35
       Val Thr Glu Phe Tyr Arg Phe Arg Gly Arg Lys Ala Met Asn Thr Gly 305 310 315
        Arg Arg Thr Ser Leu Tyr Tyr His Asp
40
        (2) INFORMATION FOR SEQ ID NO:363
              (i) SEQUENCE CHARACTERISTICS:
45
                    (A) LENGTH: 319 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
50
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
55
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature (B) LOCATION 1...319
60
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363
       Phe Thr Ser Gly Thr Ile Phe Val Arg Ile Ala Ile Ser Gly Arg Val
       Val Cys Arg Glu Val Glu Asp Asn Gly Arg Arg Glu Ala Val Arg His
20 25 30
65
       Asp Gly Met Val Val Thr Leu Leu Val Ile Val Gly Ile Val Val Val Val 35
       Val Arg Tyr Ser Leu Arg Val His Val His Lys Thr Gly Thr Val Val 50 60
70
       Ser Ala Ala Ile Phe Gly Phe Ile Leu Leu Gly Lys Thr Val Pro Cys 65 70 70
       Asp Thr Arg Asn Phe Phe Ser Ser Glu Ser Asp Glu Pro Glu Ser Arg
                                                 90
       Val Ala Thr Glu Ile Ala His Leu Cys Glu Ile Gly Phe Gln Ile His
                                             105
```

```
Ala Ser Ser Ile His Val Ala Val Arg Thr Asp Phe Gly Gln Ala Gly 115 120 125
       The His Cys Pro Met Ala Thr Asp Ala Ser Ala Thr Glu Phe Asp Arg
                                135
           130
       Ser Ala Glu Cys Ala Glu Arg Thr Ser Ala Gln Ile Asp Thr Ala Ile
145 150 155 160
       Arg Ser Gln Ser Gln Ile Ile Arg Thr His Ile Asp Thr Cys Pro Lys
165 170 175
       Ser Ser Gly Thr Ile Gly Gly Ser Thr His Thr Ser Leu His Leu Lys
180 190
10
       Val Phe Asp Gly Arg Gly Glu Val Gly His Ile His Pro Lys Asp Gly
195 200 205
       Leu Arg Phe Gly Val Val Glu Gly Tyr Ser Ile Gly Ser Tyr Val Asp
210 215 220
15
       Ala Ile Gly Ile Gly Ala Thr His Ala Lys Ala Gly Ile Ser Asp Thr
225 230 235 240
       Arg Thr Gly Ile Ala Gly Gly Tyr His Gly Gly Ser Gln His Gln Gln 245 250 255
       Ile Gly Asp Val Thr Thr Ile Ile Gly Leu Gly Lys Phe Gly Leu Ala 260 265 270
20
       Asn Val Gly Val Ser Asp Arg Gly Phe Arg Arg Gly Ala Ser Ser Tyr
275 280 285
       Asp Leu Tyr Gly Leu Glu Leu His Ile Ala Lys Thr Ile Tyr Leu Ala 290 295 300
25
       Val Gly Asp Gly Cys Leu Cys Arg Gln Ala Glu Arg Gln Asp Gly
                              310
       (2) INFORMATION FOR SEQ ID NO:364
30
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 614 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
35
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
40
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...614
45
            (Mi) SEQUENCE DESCRIPTION: SEQ ID NO:364
       Pro Tyr Asn Val Gln Ser Ile Ser Asn Lys Thr Ile Lys Lys Gln Met
1 5 10 15
50
       Glu Asn Leu Lys Asn Ile Gln Pro Arg Glu Asp Phe Asn Trp Glu Glu 20 25 30
       Phe Glu Ala Gly Gly Val His Ala Ala Val Ser Arg Gln Glu Gln Glu
35 40
       Ala Ala Tyr Asp Lys Thr Leu Asn Thr Ile Lys Glu Lys Glu Val Val 50 60
55
       Het Gly Arg Val Thr Ala Ile Asn Lys Arg Glu Val Val Ile Asn Val
65 70 75 80
       Gly Tyr Lys Ser Glu Gly Val Val Pro Ala Thr Glu Phe Arg Tyr Asn
85 90 95
       Pro Glu Leu Lys Val Gly Asp Glu Val Glu Val Tyr Ile Glu Asn Gln
100 105 110
60
       Glu Asp Lys Lys Gly Gln Leu Val Leu Ser His Arg Lys Gly Arg Ala
115 120 125
       Ala Arg Ser Trp Glu Arg Val Asn Glu Ala Leu Glu Lys Asp Glu Ile
130 135 140
65
       Val Lys Gly Tyr Val Lys Cys Arg Thr Lys Gly Gly Met Ile Val Asp
145 150 155 160
       Val Phe Gly Ile Glu Ala Phe Leu Pro Gly Ser Gln Ile Asp Val Arg
165 170
70
       Pro Ile Arg Asp Tyr Asp Ala Phe Val Glu Lys Thr Het Glu Phe Lys
180 185 190
       Ile Val Lys Ile Asn Gln Glu Tyr Lys Asn Val Val Val Ser His Lys
195 200 205
       Val Leu Ile Glu Ala Glu Leu Glu Gln Gln Lys Lys Glu Ile Ile Gly
75
```

```
Lys Leu Glu Lys Gly Gln Val Leu Glu Gly Ile Val Lys Asn Ile Thr
       Ser Tyr Gly Val Phe Ile Asp Leu Gly Gly Val Asp Gly Leu Ile His 245 250 255 Ile Thr Asp Leu Ser Trp Gly N=2 V3 \frac{1}{2}
 5
       Ile Thr Asp Leu Ser Trp Gly Arg Val Ala His Pro Glu Glu Ile Val
260 265 270
       Gln Leu Asp Gln Lys Ile Asn Val Val Ile Leu Asp Phe Asp Glu Asp
275 280 285
       Arg Lys Arg Ile Ala Leu Gly Leu Lys Gln Leu Met Pro His Pro Trp
290 295 300
10
       Asp Ala Leu Asp Ser Glu Leu Lys Val Gly Asp Lys Val Lys Gly Lys 305 310 315 320
                                                    315
       Val Val Wat Ala Asp Tyr Gly Ala Phe Val Glu Ile Ala Gln Gly
                                              330
                      325
       Val Glu Gly Leu Ile His Val Ser Glu Met Ser Trp Thr Gln His Leu
340 345 350
15
       Arg Ser Ala Gln Asp Phe Leu His Val Gly Asp Glu Val Glu Ala Val 355 360 365
       Ile Leu Thr Leu Asp Arg Glu Glu Arg Lys Met Ser Leu Gly Leu Lys 370 375 380
20
       Gln Leu Lys Pro Asp Pro Trp Ala Asp Ile Glu Thr Arg Phe Pro Val
385 390 395 400
       Gly Ser Arg His His Ala Arg Val Arg Asn Phe Thr Asn Phe Gly Val
405 415
25
       Phe Val Glu Ile Glu Glu Gly Val Asp Gly Leu Ile His Ile Ser Asp 420 425 430
       Leu Ser Trp Thr Lys Lys Ile Lys His Pro Ser Glu Phe Thr Glu Val
435 440 445
       Gly Ala Asp Ile Glu Val Gln Val Ile Glu Ile Asp Lys Glu Asn Arg
450 455 460
30
       Arg Leu Ser Leu Gly His Lys Gln Leu Glu Glu Asn Pro Trp Asp Val 465 470 475 480

Phe Glu Thr Val Phe Thr Val Gly Ser Ile His Glu Gly Thr Val Ile 485 490 490 495
35
       Glu Val Net Asp Lys Gly Ala Val Val Ser Leu Pro Tyr Gly Val Glu
500 510
       Gly Phe Ala Thr Pro Lys His Met Val Lys Glu Asp Gly Ser Gln Ala
515 520 525
       Val Leu Glu Glu Lys Leu Pro Phe Lys Val Ile Glu Phe Asn Lys Asp
530 535 540
40
                                                       540
       Ala Lys Arg Ile Ile Val Ser His Ser Arg Val Phe Glu Asp Glu Gln 545 550 560
       Lys Met Ala Gln Arg Glu Ala Asn Ala Glu Arg Lys Ala Glu Ala Lys
565 570 575
45
       Ala Ala Glu Lys Glu Ala Ala Ala Glu Ala Ala Asn Pro Ala Gln Ala
580 585 590
       Val Glu Lys Ala Thr Leu Gly Asp Leu Gly Glu Leu Ala Ala Leu Lys
             595
       Glu Lys Leu Ser Glu Asn
50
       (2) INFORMATION FOR SEQ ID NO:365
             (i) SEQUENCE CHARACTERISTICS:
55
                   (A) LENGTH: 243 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
60
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
65
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...243
70
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365
       Ser Thr Ile Met Lys Lys Ala Ile Leu Ser Gly Ala Ala Leu Leu Leu
                                                 10
       Gly Leu Cys Ala Asn Ala Gln Asn Val Gln Leu His Tyr Asp Phe Gly
75
```

```
His Ser Ile Tyr Asp Glu Leu Asp Gly Arg Pro Lys Leu Thr Thr Thr 35 40 45
       Val Glu Asn Phe Thr Pro Asp Lys Trp Gly Ser Thr Phe Phe Phe Ile
50 55 60
       Asp Met Asp Tyr Thr Gly Lys Gly Ile Gln Ser Ala Tyr Trp Glu Ile 65 70 75
       Ser Arg Glu Leu Lys Phe Trp Gln Ala Pro Val Ser Ile His Leu Glu
85 90 95
       Tyr Asn Gly Gly Leu Ser Thr Ser Phe Thr Phe Gly His Asp Ala Leu
100 105 110
10
       Ile Gly Ala Thr Tyr Thr Tyr Asn Asn Pro Ser Phe Thr Arg Gly Phe 115 120 125
       Thr Ile Thr Pro Met Tyr Lys His Leu Gly Ala His Asp Phe His Thr
130 135 140
       Tyr Gln Ile Thr Gly Thr Trp Tyr Met His Phe Leu Asp Gly Leu Leu
145 150 160
15
       Thr Phe Asn Gly Phe Leu Asp Leu Trp Gly Phe Pro Gln Glu Asn Pro
165 170 175
       Ile Gly Gly Pro Val Leu Lys Glu Gly Asp Lys Phe Val Phe Leu Ser 180 185 190

Glu Pro Gln Phe Trp Ile Asn Leu Asn Arg Ile Lys Gly Ile Asp Lys 195 200 205
20
       Asp Phe Asn Leu Ser Ile Gly Thr Glu Met Glu Ile Ser Arg Asn Phe
210 215 220
                                                         220
25
       Ala Arg Met Asp Lys Phe Ser Cys Ile Pro Thr Leu Ala Val Lys Trp
       Thr Phe Asn
30
       (2) INFORMATION FOR SEQ ID NO: 366
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 235 amino acids
                   (B) TYPE: amino acid
35
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
40
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
45
                   (A) NAME/KEY: misc_feature
(B) LOCATION 1...235
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:366
50
       Met Lys Thr Ile Ser Lys Asn His Ala Ala Arg Ile Cys Ala Ala Ile
       Ala Leu Phe Ala Val Cys Asn Gly Arg Ile Ala Ala Gln Asp Phe Leu
20 25 30
       Tyr Glu Ile Gly Gly Gly Phe Gly Ala Ala Gln Tyr Phe Gly Asp Ala 35 40
55
       Asn Arg Gly Leu Phe Gly Ser Ser Gly Val Gly Leu Glu Leu Val Gly
50 55 60
       Arg Tyr Asn Tyr Asn Phe Arg Trp Ala Phe Ser Thr Het Leu Asp Trp 65 70 75 80
60
       Arg Thr Leu Arg Gly Asp Thr Asp Lys Ser Gly Asn Val Phe Pro Asp
85 90
       Phe Ala Gln Ala Asp Phe Lys Val Gly Leu Thr Gln Leu His Val Arg 100 \hspace{1cm} 105 \hspace{1cm} 110
       Ser Glu Phe Asn Phe Leu Pro Tyr Ser Asp Gly Tyr Lys Tyr Leu Gly 115 120 120
65
       Thr Ala Arg Leu Ser Pro Tyr Val Ala Ala Gly Leu Ser Leu Gly Phe
130 135 140
       Ala Ser Gly Ala Lys Gly Ser Ala Phe Ala Pro Gly Ile Thr Ala Gly 145 150 150
       Met Gly Val Lys Tyr Lys Leu Lys Pro Arg Ile Asn Val Gly Ile Glu
165 170 175
70
       Tyr Ser Phe Thr Gly Leu Leu Thr Asp Ala Leu Asp Ala Leu Thr Asp
180 185 190
       Lys Ser Val Trp Leu Glu Asp Pro Tyr Lys Ile Asn Asp Ser Trp Val
75
                                        200
```

```
Lys Asn Lys Asp Ala Thr Gly Ala Leu Val Leu Arg Ile Thr Tyr Asp 210 215 220
        Phe Gly Leu Arg Lys Thr Phe Cys Asn Lys Gln
  5
        (2) INFORMATION FOR SEQ ID NO:367
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 436 amino acids
 10
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
 15
           (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
 20
             (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...436
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367
25
        Asn Ile Met Tyr Lys Asp Tyr Lys Gly Leu Tyr Ala Ser Leu Arg Trp
                                                  10
        Tyr Ala Leu Ile Ile Gly Leu Leu Phe Ala Ala Asp Gly Ile Gln Ala 20 25 30
30
        Gln Asn Asn Phe Thr Glu Ser Pro Tyr Thr Arg Phe Gly Leu Gly
35 40
        Arg Leu Gly Glu Arg Thr Thr Ile Ser Gly His Ser Met Gly Gly Leu 50 60
        Gly Val Gly Leu Arg Gln Gly Thr Tyr Val Asn Ala Val Asn Pro Ala
65 70 75
35
       Ser Tyr Ser Ala Val Asp Ser Met Thr Phe Ile Phe Asp Phe Gly Ala
85 90
       Ser Thr Gly Ile Thr Trp Tyr Ala Glu Asn Gly Lys Lys Asp Asn Arg
100 105
       Lys Net Gly Asn Ile Glu Tyr Phe Ala Met Leu Phe Pro Ile Ser Lys
40
       Ser Ile Ala Met Ser Ala Gly Val Leu Pro Tyr Ser Ala Ser Gly Tyr
130 135 140
       Gln Fhe Gly Ser Val Asp Gln Val Glu Gly Gly Ser Val Gln Tyr Thr
145 150 155 160
45
       Arg Lys Tyr Leu Gly Thr Gly Asn Leu Asn Asp Leu Tyr Val Gly Ile

165 170 175

Gly Ala Thr Pro Phe Lys Asn Phe Ser Ile Gly Ala Asn Ala Ser Ser

180 185 190
       Leu Phe Gly Arg Phe Thr His Ser Arg Gln Val Ile Phe Ser Thr Glu
195 200 205
50
       Ala Pro Tyr Asn Pro Val His Leu Ser Thr Leu Tyr Leu Lys Ala Ala
210 215 220
       Lys Phe Asp Phe Gly Met Gln Tyr His Leu Leu Lys Ser Asp Arg 225 230 235 240
55
       Ser Leu Val Ile Gly Ala Val Tyr Ser Pro Arg Val Lys Met His Ser
245 250 255
       Glu Leu Thr Gln Ile Lys Asn Gln Val Gln Asn Gly Val Val Val Glu
260 265 270
60
       Ser Glu Thr Gln Glu Tyr Ile Lys Gly Met Asp Tyr Tyr Thr Leu Pro
275 280 285
       His Thr Leu Gly Ile Gly Phe Ser Tyr Glu Lys Lys Asp Lys Leu Leu
290 295 300
       Leu Gly Ala Asp Val Gln Tyr Ser Lys Trp Lys Gly Glu Lys Phe Tyr 305 310 315
65
       Lys Ser Asp Cys Lys Phe Gln Asp Arg Ile Arg Val Ser Leu Gly Gly 325 330 335
       Glu Ile Ile Pro Asp Ile Asn Ala Val Gly Net Trp Pro Lys Val Arg
340 345 350
70
       Tyr Arg Phe Gly Leu His Gly Glu Asn Ser Tyr Leu Lys Val Pro Thr
355 360 365
       Lys Gly Gly Val Tyr Gln Gly Tyr His Ile Val Gly Ala Val Phe Gly 370 375 380
       Ile Gly Ile Pro Leu Asn Asp Arg Arg Ser Phe Val Asn Val Ser Leu
385 390 395 400
75
```

```
Glu Tyr Asp Arg Leu Ile Pro Lys Glu Gly Met Ile Lys Glu Asn Ala
405 410
                         405
                                              410
        Leu Lys Leu Thr Phe Gly Leu Thr Phe Asn Glu Ser Trp Phe Lys Lys
                     420
  5
        Leu Lys Leu Asn
                 435
        (2) INFORMATION FOR SEQ ID NO:368
 10
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 945 amino acids
(B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 15
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
 20
                   (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...945
25
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:368
        Asn Thr Ile Ser Glu Asn Gly Asn Asp Ser Thr Tyr Phe Ser Phe Gln
30
        Thr Phe Leu Met Arg Ser Ile Tyr Gln Leu Leu Leu Ser Ile Leu Leu
                    20
                                            25
                                                                  30
        Ala Ser Leu Gly Phe Val Gly Leu Glu Ala Gln Gln Ala Gly Val Ala
                                       40
                                                             45
       Gly Arg Val Leu Asp Glu Glu Gly Asn Pro Het Ile Gln Ala Asn Val
50 55 60
35
       Gln Leu Val Gln Ser Thr Gly Gln Val Ala Val Ala Ala Gly Ala Thr
65 70 75 80
       Asn Glu Lys Gly Leu Phe Ser Leu Lys Thr Ser Gln Glu Gly Asp Tyr
85 90 95
40
       Ile Leu Arg Val Ser Tyr Val Gly Tyr Thr Thr His Asp Glu Lys Ile 100 $105\ 
       Ser Leu Arg Asn Gly Gln Thr Ile Thr Leu Lys Asp Ile Ser Met Asn
115 120 125
       Glu Asp Ala Arg Leu Leu Gln Ser Val Thr Val Gln Ala Lys Ala Ala
130 135 140
45
       Glu Val Val Val Arg Asn Asp Thr Leu Glu Phe Asn Ala Gly Ser Tyr
145 150 155 160
       Thr Val Ala Gln Gly Ala Ser Ile Glu Glu Leu Ile Lys Lys Leu Pro
165 170 175
50
       Gly Ala Glu Ile Gly Ser Asp Gly Lys Ile Thr Ile Asn Gly Lys Asp 180 185 190
       Ile Ser Lys Ile Leu Val Asp Gly Lys Glu Phe Phe Ser Lys Asp Pro
195 200 205
       Gln Val Ala Ile Lys Asn Leu Pro Ala Asp Met Val Asn Lys Val Gln
210 215 220
55
       Val Leu Asn Lys Leu Ser Glu Leu Ser Arg Het Ser Gly Phe Asp Asp 225 230 235 240
       Gly Glu Glu Glu Thr Val Ile Asn Leu Thr Val Lys Pro Glu Lys Lys 245 250 255
60
       Lys Gly Leu Phe Gly Thr Leu Gln Ala Gly Tyr Gly Thr Asp Gln Arg 260 265 270
       Tyr Met Ala Gly Gly Asn Val Asn Arg Phe Asp Gly Asn Lys Gln Trp 275 280 285
       Thr Leu Ile Gly Ser Ala Asn Asn Thr Asn Asn Het Gly Phe Ser Glu
290 295 300
65
                                                       300
       Het Asp Ser Glu Met Gly Ser Met Thr Phe Phe Ser Pro Gln Gly Gly 305 310 315 320
       305 310 315 320 Gly Arg Arg Gly Phe Gly Asn Ser Gly Gly Val Thr Ser Ser Met 325 330 335
70
       Leu Gly Gly Asn Phe Ser Val Glu Phe Ser Ser Ala Leu Asn Thr Gly 340 345
                                          345
                                                               350
       Gly Asp Ala Arg Tyr Gly Tyr Asn Asp Lys Ala Ile Glu Thr Thr Lys
355 360 365
       Arg Val Glu Asn Ile Leu Ala Glu Gly Asn Thr Tyr Met Asp Glu Asn
75
                                                         380
```

```
Ile Leu Glu Arg Ser Phe Ser His Asn Gly Gln Ala Arg Phe Arg Met 385 390 395 400
         Gln Trp Lys Pro Ser Glu Arg Thr Glu Val Val Phe Glu Pro Asp Leu
405
410
415
         Ser Ile Ser Lys Ile Asp Gly Phe Phe Asn Asp Thr Tyr Glu Thr Lys
420 425 430
  5
         Asp Ala Thr Gly Ile Ser Ile Asn Lys Gly Ser Ile His Gln Thr Thr 435 440 445
         Gln Gly Asn Asn Phe Arg Leu Asn Gly Glu Leu Asp Ile Ser His Lys
450 455 460
 10
         Leu Asn Asp Glu Gly Arg Thr Ile Ser Ala Ser Val Ser Gly Gly Leu
465 470 475
         Thr Asp Glu Asp Gly Asp Gly Ile Tyr Gln Ala Val Leu Gln Ser Val
485 490 495
         Glu Thr Asn Gln Lys Gln Phe Asn Asp Asn Ser Asn Leu Gln Tyr Arg
500 505 510
 15
         Leu Arg Leu Ser Tyr Val Glu Pro Leu Gly Lys Asn Tyr Phe Ala Gln
515 520 525
         Ala Ile Leu Asn Arg Arg Phe Ser Arg Arg Asn Ser Asp Arg Glu Val
530 535 540
 20
         Tyr Arg Leu Gly Asp Asp Gly Gln Tyr Ser 11e Leu Asp Ser Gln Tyr
545 550 555 560
        Gly Leu Ser Tyr Ser Asn Glu Phe Thr Gln Tyr Arg Ile Gly Leu Asn
565 570 575
 25
         Leu Lys Lys Ile Ala Lys Thr Trp Asp Tyr Thr Val Gly Phe Asn Val 580 585
        Asp Pro Asn Arg Thr Val Ser Tyr Arg Ser Val Ala Gly Val Glu Gln 595 600 605
        Asp Lys Leu Ala Phe Asn Arg Val Asn Leu Ser Pro Met Leu Arg Ile
610 615 620
30
        Asn Tyr Lys Pro Ser Arg Thr Thr Asn Leu Arg Val Asp Tyr Arg Gly 625 630 635 640
        Arg Thr Thr Gln Pro Ser Ile Asn Gln Ile Ala Pro Val Gln Asp Ile
645
650
655

Thr Asn Pro Leu Phe Val Thr Glu Gly Asn Pro Gly Leu Lys Pro Ser
660
665
670
35
        Tyr Ser Asn Asn Val Met Ala Met Phe Ser Asp Phe Asp Ala Lys Ser
675 680 685
        Gin Arg Ala Phe Asn Ile Val Phe Phe Gly Asn Tyr Thr Phe Asp Asp 690 695 700

Ile Val Pro Asn Thr His Tyr Asp Pro Ser Thr Gly Ile Arg Thr Thr 705 710 720
40
        Arg Tyr Glu Asn Ala Ser Gly Thr Trp Gln Ala Asn Leu His Gly Thr
725 730 735
        Leu Ser Leu Pro Leu Lys Asn Arg Ala Phe Ser Phe Arg Met Ser Leu 740 745 750
45
        Phe Asn Arg Leu Ala Glu Gly Gln Ser Phe Ile Asn Asp Asp Lys Asn 755 760 765
        Lys Ala Leu Ser Phe Arg Thr Arg Glu Arg Leu Thr Leu Thr Tyr Arg
770 775 780
50
        Asn Asn Trp Ile Asp Thr Ser Ile Gly Gly Asn Ile Gly Phe Tyr Net
785 790 795 800
        Ala Asn Asn Ser Leu Ser Gly Gln Lys Asp Ser Arg Thr Tyr Asp Phe
805 810 815
55
        Gly Gly Asn Tyr Gln Val Ala Leu Thr Leu Pro Tyr Gly Phe Arg Ile
820 825 830
        Asp Ser Asp Val Glu Tyr Asn Thr Asn Ser Gly Tyr Ser Gly Gly Phe
835 840 845
        Ser Leu Asp Glu Trp Leu Trp Asn Ala Ser Leu Ser Tyr Ser Phe Leu 850 855 860
60
        Arg Asp Lys Ala Gly Thr Leu Arg Val Asn Gly Tyr Asp Ile Leu Gly 865 870 875 880 Gln Arg Ser Ser Ile Ser Arg Ser Ala Ser Ala Ile Asn Ile Glu Glu 895 895
65
        Ser Het Ser Asn Thr Ile Gly Arg Tyr Val Met Val Asp Phe Ile Tyr 900 905 910
        Arg Phe Asn Ala Phe Ser Gly Gly Gly Ser Arg Ser Asp His Gln Arg
915 920 925
        Gly Asn Net Asn Arg Pro Gly Pro Pro Phe Gly Gly Gly Arg Arg Pro
930 935 940
70
        Ser
        945
        (2) INFORMATION FOR SEQ ID NO:369
75
```

```
(i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 412 amino acids
                     (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
 5
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
10
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
15
                    (B) LOCATION 1...412
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:369
       Gly Glu Tyr Pro Ala Asn Ser Asn Asp Lys Lys Glu Met Val Het Lys 1 \phantom{\bigg|} 5 \phantom{\bigg|} 10 \phantom{\bigg|} 15
20
       Leu Ile Lys Arg Ser Leu Leu Leu Gly Ala Val Leu Leu Ile Thr
20 25 30
       Leu Pro Ala Tyr Ser Gln Asn Asp Asp Ile Phe Glu Asp Asp Ile Tyr 35 40 45
25
       Thr Ser Arg Lys Glu Ile Arg Lys Gln Asn Gln Val Lys Asp Trp Gln 50 60
       Asn Gln Glu Asp Gly Tyr Gly Asp Asp Thr Glu Tyr Thr Val Ala Ser 65 70 75 80
       Asp Arg Asp Ile Asp Ala Tyr Asn Arg Arg Asp Gly Gln Ser Tyr Asp 85 90 95
30
       Gly Lys Lys Leu Ser Lys Asp Lys Lys Arg Asp Ser Thr Arg Ser Ser 100 105 110
       Val Pro Gly Arg Tyr Ser Arg Arg Leu Ala Arg Phe Tyr Lys Pro Asn
115 120 125
35
       Thr Ile Val Ile Ser Gly Ala Asp Asn Val Tyr Val Thr Asp Asp Gly 130 135 140
       Glu Tyr Phe Val Tyr Gly Asp Glu Tyr Tyr Asp Asp Ala Ser Ser Val
145 150 155 160
       Asn Ile Tyr Ile Asn Ser Pro Trp Cys Asp Pro Phe Pro Tyr Thr Ser
165 170 175
40
       Trp Tyr Pro Ser Phe Ser Gly Trp Tyr Asn Tyr Thr Trp Asn Tyr Pro
180 185 190
       Trp Phe Tyr Tyr Gly Ser His Ile Gly Trp Gly Gly Tyr Tyr Pro Gly 195 200 205
45
       Tyr Asn Trp Tyr Trp Ser Tyr Tyr Tyr Asp Pro Phe Tyr Asn Pro Tyr 210 215 220
       Gly Ile Gly Met Gly Trp Gly Tyr Pro Tyr Gly Trp Gly Ser Tyr Tyr 225 230 235 240 Gly Trp Gly Gly Tyr Pro Gly Val Ile His His Tyr His His Tyr Pro 245 250 255
50
       Lys Lys Thr Tyr Ser Asn Gly Gln His Ser Gly Ala Tyr Tyr Ser Tyr
260 265 270
       Gly Arg Pro Asn Arg Ile Lys Gly Gly Thr Ser Gly Ala Lys Leu Gly
275 280 285
55
       Thr Gly Arg Tyr Asp Arg Ile Gln Asn Ser Ser Ser Gln Lys Asn Lys
290 295 300
        Phe Gly Leu Gln Ser Asn Lys Pro Asn Asn Asn Leu Gln Asn Val Lys
305 310 315 320
       Ser Gly Arg Thr Gly Arg Ala Asn Arg Asp Arg Asn Ile Glu Thr Val
325 330 335
60
       Thr Pro Asn Asn Gly Gln Lys Gln Asn Arg Pro Val Phe Gln Gln Asn 340 345 350

Gln Ser Gly Asn Asp Arg Pro Thr Gly Arg Asn Ile Arg Ser Glu Arg 355 360 365
65
       Gln Gly Glu Asn Asn Asp Arg Thr Phe Ser Thr Pro Ser Arg Ser Asn 370 380
       Ser Asn Gly Gly Phe Ser Thr Pro Ser Arg Ser Ser Ser Gly Ser Met
385 390 395 400
        Ser Gly Gly Gly Arg Ser Gly Arg Gly Arg Asn
70
        (2) INFORMATION FOR SEQ ID NO:370
              (i) SEQUENCE CHARACTERISTICS:
75
                    (A) LENGTH: 601 amino acids
```

```
(B) TYPE: amino acid
                      (D) TOPOLOGY: linear
              (ii) MOLECULE TYPE: protein
  5
             (iii) HYPOTHETICAL: YES
              (vi) ORIGINAL SOURCE:
                      (A) ORGANISM: Porphyromonas gingivalis
 10
              (ix) FEATURE:
                      (A) NAME/KEY: misc feature
                      (B) LOCATION 1...601
15
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:370
        Ser Asn Ser Ser Ser His Lys Trp Leu Ile Tyr Tyr His Ile Glu Lys 1 \phantom{\bigg|} 5 \phantom{\bigg|} 10 \phantom{\bigg|} 15
        Thr Lys Ser Ile !let Ile Arg Lys Leu Ile Leu Leu Leu Ala Leu Met 20 \hspace{1cm} 25 \hspace{1cm} 30
20
        Pro Val Ala Ser Val Ala Phe Ala Val Pro Thr Asp Ser Thr Glu Ser 35 40 45
        Lys Asp Asn Arg Ile Leu Thr Ser Met Gln Ser Ser Leu Asn Arg
50 55 60
25
        Asp Asp Ala Pro Asp Lys Trp Gln Pro Het His Ala Asn Phe Ser IIe 65 70 75 80
        Gln Ser Asp Het Leu Leu Ser Thr Ala Gln Lys Ser Lys Asn Thr Trp
85 90 95
Phe Gly Asn Ser Tyr Ile Het Gly Ile Ile Lys Asn Asn Tyr Leu Glu
100 105 110
30
        Phe Gly Ala Arg Phe Glu Asp Leu Tyr Lys Pro Leu Pro Gly His Glu 115 120 125

Pro Glu Met Gly Arg Gly Val Pro His Met Tyr Val Lys Gly Ser Tyr 130 135 140
        His Trp Ala Glu Leu Thr Met Gly Asp Phe Tyr Asp Gln Phe Gly Ser
35
        Gly Net Val Phe Arg Thr Tyr Glu Glu Arg Asn Leu Gly Ile Asp Asn
165 170 175
        Ala Val Arg Gly Gly Arg Ile Val Leu Thr Pro Phe Asp Gly Val Arg
180 185
40
        Val Lys Gly Ile Ala Gly Gln Gln Arg Asn Tyr Phe Asp Arg Thr Gly
195 200 205
        Lys Val Phe Asn Ser Gly Arg Gly Tyr Leu Leu Gly Ser Asp Leu Glu
210 215 220
        Leu Asn Val Glu Arg Trp Ser Ser Ala Net Arg Asp Asn Asp Tyr His
225 230 235 240
45
        Leu Ala Ile Gly Gly Ser Phe Val Ser Lys His Glu Ala Asp Glu Asp
245
250
255

Ile Phe Val Gly Val Gly Glu Asp Arg Lys Arg Leu Asn Leu Pro Leu
260
265
270
50
        Asn Val Pro Ile Het Gly Leu Arg Thr Asn Phe Gln Lys Gly Gly Leu
275 280 285
        Ala Leu Tyr Ala Glu Tyr Gly Tyr Lys Tyr Asn Asp Pro Ser Ala Asp
290 295 300
55
       Asn Asp Tyr Ile Tyr His Asp Gly Gln Ala Ala Leu Leu Ser Ala Ser 305 310 315 320
       Tyr Ser Lys Lys Gly Met Ser Ile Leu Leu Gln Ala Lys Arg Cys Glu
325
Asn Phe Ala Phe Arg Ser Lys Arg Ser Ala Gln Leu Thr Pro Leu Met
340
345
350
60
        Ile Asn Tyr Met Pro Ala Phe Thr Gln Ala His Thr Tyr Thr Leu Ala
355 360 365
       Ala Ile Tyr Pro Tyr Ala Thr Gln Pro Gln Gly Glu Trp Ala Phe Gln 370 375 380
65
       Gly Glu Leu Arg Tyr Asn Phe Ala Arg Arg Thr Ala Leu Gly Gly Arg
385 390 395
       Tyr Gly Thr Gly Leu Arg Ile Asn Val Ser His Val Arg Gly Leu Asp
405 415
       Lys Lys Het Leu Lys Glu Asn Pro Asp Glu Leu Ile Gly Thr Asp Gly 420 425 430
70
       Tyr Thr Val Ser Phe Phe Gly Het Gly Asp Leu Tyr Tyr Ser Asp Ile
435 440
       Asp Val Glu Ile Thr Lys Lys Val Ser Pro Gly Phe Asp Phe Thr Leu
450 455 460
75
        Thr Tyr Leu Asn Gln Ile Tyr Asn Asn Lys Val Leu His Gly Ala Ala
```



```
Gly Glu Lys Pro Glu Lys Ile Tyr Ala Asn Ile Phe Val Tyr Asp Gly
485 490 495
       Lys Tyr Lys Leu Ser Asn Lys Val Ala Leu Arg Thr Glu Leu Gln Tyr 500 505 510
                                                               510
       Leu His Thr Lys Gln Asp Gln Gly Asp Trp Ile Tyr Gly Met Ala Glu
515 520 525
       Leu Ser Ile Leu Pro Ser Leu Met Leu Ser Leu Ser Glu Gln Tyr Asn
530 535 540
10
       Ile Gly Glu Thr Lys Lys His Tyr Val Met Gly Ser Val Thr Tyr Thr 545 550 560
       His Gly Ala His Arg Val Ala Phe Ser Ala Gly Lys Thr Arg Ala Gly
565 570 575
       Het Asn Cys Ser Gly Gly Val Cys Arg Val Val Pro Glu Thr Gln Gly 580 585
15
       Phe Tyr Leu Ser Tyr Ser Thr Asn Leu
       (2) INFORMATION FOR SEQ ID NO:371
20
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 252 amino acids
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
           (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
35
                  (B) LOCATION 1...252
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:371
       Arg Gly Ser Ser Ser Gly Ile Ser Ala Arg Gly Arg Asp Met Arg Ser 1 5 10 15
40
       Leu Phe Leu Ser Ala Leu Arg Ser Ser Ser Leu His Gly Ser Glu Arg 20 25 30
       Arg Ser Arg Ile Ser Ser Ser Val Val Met Ser Ile Arg Gln Lys Ile
35 40 45
       Arg Leu Phe His Leu Ser Val Cys Ala Gln Thr His Asp His Leu Ile
50 55 60
45
       Glu Ile His Leu Val Cys Ile Glu Phe Gly Ala Ile Asp Thr Asp Glu 65 70 75 80
       Phe Arg Leu Ser Ser His Ala Tyr Thr Thr Ser Pro Thr His Thr Gly 85 \phantom{\bigg|}90\phantom{\bigg|}
50
       Ala Ile His His Asn Cys Ile Glu Arg Ser Tyr Gly Arg Tyr Leu Val
       Thr Phe Gly Glu Arg Asn Glu Leu His His His Ser Arg Pro Asp
115 120 125
55
       Arg Asn Ala Glu Val Tyr Arg Phe Pro Phe Asp Asn Ala Phe His Ser
130 135 140
       Ile Arg Tyr Glu Ala Phe Arg Pro Ile Arg Pro Ile Val Cys His Asp
145 150 155 160
       145
       Asp His Phe Ile Ala Ile Gly Ser His Leu Phe Phe Lys Asp Asn Gln
165 170 175
60
       Ile Phe Ser Ser Gly Ser Gln Tyr Asp Asn Tyr Thr Val Ala Cys Phe
180 185 190
       Val Glu Ser Leu His Asp Arg Glu Gln Arg Ser His Thr His Thr Ala
195 200 205
       Ser Gly Thr Asn His Cys Ala Asp Leu Phe Asp Met Arg Thr Leu Ser
210 215 220
                                 215
                                                       220
       Gln Arg Thr Tyr His Ile Arg Asp Ile Val Ala Asp Phe Glu Phe Gly
                           230
       Gln Phe Leu Gly Arg Phe Ala His Gly Leu Asn His
70
       (2) INFORMATION FOR SEQ ID NO: 372
             (i) SEQUENCE CHARACTERISTICS:
75
```

(A) LENGTH: 790 amino acids

| | | | | B) T | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| _ | | (ii |) 110 | LECU: | LE T | PE: | pro | tein | | | | | | | | |
| 5 | | (iii |) HY: | РОТНІ | ETIC | AL: | YES | | | | | | | | | |
| | | (vi |) OR: | IGIN | AL S | OURC | E: | | | | | | | | | |
| 10 | | , | • | | | | | phyro | nomo | as g | ingi | vali | S | | | |
| | | (ix | | A) N | AI IE/I | | | c_fea | atur | e | | | | | | |
| 4= | | | (i | B) L | OCAT: | 1101 | 1 | 790 | | | | | | | | |
| 15 | | (xi |) SE(| DOEH | CE DE | ESCR. | I PTI | 014: 3 | SEQ : | ID N | 0:37 | 2 | | | | |
| | Leu 1 | Leu | Phe | Ser | Ser 5 | Pro | Leu | Pro | Arg | Met 10 | Asp | Arg | Pro | Lys | Pro 15 | Ser |
| 20 | Tyr | Ile | Val | Arg 20 | Ile | Ala | Ala | Ile | Leu 25 | | Leu | Phe | Val | Gly 30 | Arg | Pro |
| | Leu | Phe | Ala 35 | | Ser | Tyr | Val | 40 Asp | | Val | Asp | Pro | Leu 45 | | Gly | Thr |
| | Leu | Ser 50 | Ser | Phe | Glu | Leu | Ser 55 | | Gly | Asn | Thr | | | Val | Ile | Gly |
| 25 | | | Trp | Gly | Het | | | Trp | Thr | Pro | | 60 Thr | Gly | Val | Pro | |
| | 65 Asp | Gly | Trp | Gln | | 70 Thr | Tyr | Ser | Ala | | 75 Lys | Ile | Arg | Gly | | Lys 80 |
| 20 | Gln | Thr | His | | 85 Pro | Ser | Pro | Trp | | 90 Asn | Asp | Tyr | Gly | Gln | 95 Phe | Ser |
| 30 | Leu | Leu | Pro | 100 Leu | Thr | Ala | Pro | | 105 Lys | Pro | Ser | Ser | | 110 Asp | Ser | Ile |
| | Ala | Leu | 115 Thr | Lys | Trp | Суѕ | Lys | 120 Gln | Leu | Phe | Ser | Asp | 125 Glu | Gln | Thr | Ser |
| 35 | Trp | 130 Phe | Ser | His | Lys | Ala | 135 Glu | Thr | Ala | Thr | Pro | 140 Tyr | Tyr | Tyr | Ser | Val |
| | 145 | | Ala | | | 150 | | | | | 155 | | | | | 160 |
| | | | Ile | | 165 | | | | | 170 | | | | | 175 | - |
| 40 | | | Trp | 180 | | | | | 185 | | | | | 190 | | |
| | | | 195 Asp | | | | | 200 | | | | | 205 | | | |
| 45 | | 210 | Pro | | | | 215 | | | | | 220 | | | | |
| 10 | 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| | | | Ala | | 245 | | | | | 250 | | | | | 255 | |
| 50 | | | Arg | 260 | | | | | 265 | | | | | 270 | | |
| | | | Val 275 | | | | | 280 | | | | | 285 | | | |
| | | 290 | Glu | | | | 295 | | | | | 300 | | | | |
| 55 | Arg 305 | Glu | Ala | Trp | Asn | Lys 310 | Val | Leu | Gly | Arg | Ile 315 | His | Val | G1 u | Glγ | Glγ 320 |
| | Thr | Lys | Asp | Glu | Arg 325 | Thr | Thr | Phe | Tyr | Ser 330 | | Leu | Tyr | Arg | Cys 335 | Leu |
| 60 | Leu | Phe | Pro | Arg 340 | Arg | Phe | Tyr | Glu | Glu 345 | | Ala | Ser | Gly | Asn 350 | Phe | Val |
| | His | Tyr | Ser 355 | | Tyr | Asn | Glу | Glu 360 | | Leu | Pro | Gly | Tyr 365 | Leu | Tyr | Thr |
| | Asp | Thr 370 | G17. | Phe | Trp | Asp | Thr 375 | | Arg | Ala | Leu | Phe 380 | | Leu | Leu | Asn |
| 65 | Leu 385 | | Tyr | Pro | Asp | Glu | | Ile | Lys | Ile | | | Gly | Leu | Leu | |
| | | Tyr | Arg | Glu | | 390 Gly | Ph∈ | Phe | Pro | | 395 Trp | Ala | Ser | Pro | | 400 His |
| 70 | Arg | Asp | Сле | | 405 Ile | Gly | Asn | Asn | | 410 Ala | Ser | Val | Leu | | 415 Asp | Ala |
| , 0 | Tyr | Leu | Lys | 420 Gly | Val | Arg | Val | | 425 Asp | Thr | Arg | Thr | | 430 Net | Asn | Gly |
| | Leu | Leu | 435 His | Ala | Thr | Lys | Ala | 440 Val | His | Pro | Lys | Ile | 445 Ser | Ser | Thr | Gly |
| 7 5 | | 450 | Gl y | | | | 455 | | | | | 460 | | | | |
| | | | | | | | | | | | | | | | | |

| | 465 | | | | | 470 | | | | | 475 | | | | | 480 |
|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | | Glγ | Ile | Asp | Glu 485 | | Ala | Ala | Arg | Thr 490 | | Glu | Tyr | Ala | Tyr 495 | |
| 5 | Asp | Trp | Сув | Ile 500 | | Arg | Leu | Gly | Arg 505 | | Leu | Gly | Trp | Asp 510 | | Ala |
| | Ala | Leu | Asp 515 | | Leu | Ala | His | Arg 520 | Ser | Het | Asn | Tyr | Arg 525 | | Leu | Phe |
| | Asp | Pro 530 | Glu | Thr | Lys | Leu | Met 535 | Arg | Gly | Arg | Asn | Gln 540 | Asp | Gly | Ser | Phe |
| 10 | Arg 545 | Thr | Pro | Phe | Ser | Pro 550 | Phe | Lys | Trp | Glγ | Asp 555 | Val | Phe | Thr | Glu | Gly 560 |
| | Asn | Ala | Trp | His | Tyr 565 | Thr | Trp | Ser | Val | Phe 570 | His | Asp | Val | Gln | Gly 575 | Leu |
| 15 | Ile | Asp | Leu | Met 580 | Gly | Gly | Asp | Arg | Pro 585 | Phe | Val | Ser | Met | Leu 590 | Asp | Ser |
| | Val | Phe | Asn 595 | Thr | Pro | Pro | Met | Phe 600 | Asp | Glu | Ser | T;r | Tyr 605 | Gly | Phe | Val |
| | Ile | His 610 | G] u | Ile | Arg | Glu | Met 615 | Gln | Ile | Ala | Asp | Het 620 | Gly | Asn | Tyr | Ala |
| 20 | His 625 | Glу | Asn | G1n | Pro | Ile 630 | | His | Met | Ile | Tyr 635 | Leu | Tyr | Asn | His | Ala 640 |
| | | His | Pro | Trp | Lys 645 | | Gln | Glu | Arg | Leu 650 | | Glu | Val. | Met | Glγ 655 | Arg |
| 25 | Leu | Tyr | Arg | Pro 660 | Thr | Pro | Asp | Glγ | Tyr 665 | Суз | Glγ | Asp | Glu | Asp 670 | Asn | Glγ |
| | Gln | Thr | Ser 675 | Ala | Trp | Tyr | Vāl | Phe 680 | Ser | Ala | Leu | G1? | Phe 685 | Tyr | Pro | Val |
| | Thr | Pro 690 | Alā | Thr | Asp | Gln | Tyr 695 | Val | Leu | Gly | Ser | Pro 700 | Ile | Phe | Ser | Lys |
| 30 | Val 705 | Ile | Leu | Ser | Phe | Pro 710 | Asp | Gly | His | Lγε | Thr 715 | Val | Leu | His | Ala | Pro 720 |
| | Ala | Asn | Ser | Ala | Asp 725 | Thr | Pro | Tyr | Ile | Arg 730 | Ser | Ile | Ser | Va1 | Glu 735 | Gly |
| 35 | Lys | Glu | Trp | Ser 740 | Суѕ | Asn | Tyr | Leu | Thr 745 | His | Glu | Gln | Leu | Arg 750 | Ser | Ser |
| | Ala | Ser | Ile 755 | Gln | Trp | Met | Met | Asp 760 | Thr | Lys | Pro | Asn | Tyr 765 | Asn | Arg | Glγ |
| | Met | Lуε 770 | Glu | Ser | Asp | Arg | Pro 775 | Туr | Ser | Phe | Ser | Thr 780 | Glu | Gln | Gln | Arg |
| 40 | Arg 785 | Ala | Asn | His | Ser | Asn 790 | | | | | | | | | | |
| | (2) | INFO | ORMA' | l'ION | FOR | SEQ | ID 1 | 10:3 | 73 | | | | | | | |
| 45 | | (i) | | | | | CTER | | | | | | | | | |
| | | | (I | 3) T | PE: | ami | no ad | cid | acio | is | | | | | | |
| 50 | | , | | | | | line | | | | | | | | | |
| 30 | | | | | | | prot | tein | | | | | | | | |
| | | (iii) | | | | | | | | | | | | | | |
| 55 | | (vi) |) OR: | | | | | ohyr | omona | as g | ingi | valis | 3 | | | |
| | | (ix) |) FEA | | | | | | | | | | | | | |
| 60 | | | | | | | m150 | | atur | 3 | | | | | | |
| 00 | | (xi) |) SE | QUENC | CE DI | ESCR: | IPTI(| OH: 1 | SEQ : | ID N | o:37 | 3 | | | | |
| | Ile | Суз | Gly | Ser | Lys | Met | Asn | Leu | Ser | | Leu | Gln | Ser | Phe | | Het |
| 65 | 1 Met | Lys | Ser | | 5 Arg | Ser | Va 1 | Leu | Leu | 10 Leu | Leu | Ehe | Pro | | 15 Ser | Leu |
| | Ile | Thr | Ala 35 | 20 Leu | Gly | Cys | Ser | | 25 Asn | ГЛЗ | Ala | Ala | Glu 45 | 30 Ser | Lys | Ser |
| 70 | Val | Ser 50 | | Asp | Ser | Ala | Tyr 55 | 40 Leu | Glu | Arg | Tyr | | | Leu | Arg | Ala |
| , 5 | Asp 65 | | Asp | Thr | Pro | Ser 70 | | His | Val | Met | | 60 Ser | Tyr | Val | Tyr | |
| | | Gly | Asp | Asp | | - | Thr | Glu | Ile | | 75 Asn | ۲. Gl | Leu | Leu | | 80 Gly |
| 75 | Asp | Ser | Leu | Het | 85 Asp | Ser | Ser | Ser | Pro | 90 Glu | Asn | Ala | Met | Glu | 95 Gly | Tyr |

```
105
       Ala Gln Net Leu Gly Glu Asp Tyr Arg Ser Asn Asn Ala Glu Ala Asn
115 120 125
       Leu Gln Gly Leu Pro Ser Asp Leu Leu Asp Tyr Ile Tyr Lys Gln Glu
130 135 140
       Asn Thr Ile Ala Tyr Cys Asp Thr Gly Leu Ile Ser Thr Arg Ile Asn 145 150 155 160
       Thr Tyr Thr Tyr Glu Gly Gly Ala His Thr Glu Asn Thr Val Arg Phe
165 170 175
10
       Ala Asn Ile Leu Arg Thr Thr Gly Lys Val Leu Glu Glu Arg Asp Ile
180 185 190
       Phe Lys Ile Asp Tyr Ala Glu Arg Leu Ser Ala Leu Ile Ile Gly Gln
195 200 205
       Leu Val His Asp Phe Gly Lys Thr Thr Pro Ala Glu Leu Asp Ala Ile
210 215 220
15
       Gly Phe Phe Asn Ala Glu Glu Ile Gln Pro Asn Gly Asn Phe Met Ile
225 230 235 240
       Asp Asp Lys Gly Leu Thr Tyr Cys Phe Asn Glu Tyr Gln Ile Ala Ala 245 250 255

Tyr Ala Arg Gly Ala Val Tyr Val Arg Leu Gly Tyr Asp Val Leu Ala 260 265 270
20
       Pro Leu Leu Arg Asp Asp Ser Pro Leu Lys Arg Tyr Leu Pro
25
       (2) INFORNATION FOR SEQ ID NO: 374
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 378 amino acids
                   (B) TYPE: amino acid
30
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
35
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
(B) LOCATION 1...378
40
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:374
45
       Gly Ile Ile Ser Val Leu Ser His Val Val Asp Arg Pro Gln Trp Gly
                                               10
       Ala Ser Pro Glu Ala Ala Gly Thr His Ser Val Tyr Ser Ile Leu His
20 25 30
       Pro Ser Ala Gly Ile Ile Arg Ile Arg Ser Met Gly Ile Ile Ser Ala 35 40 45
50
       Cys Arg Ile Ala Ile Leu Ala Gly Lys His Pro Arg Ser Gly Leu Ser 50 55 60
       Arg Ala Asn Val Gly Ile Leu Ser Tyr Asn Pro Glu Asn Thr Pro Glu
65 70 75 80
       Lys Lys Arg Lys Leu Gln Glu Lys Asn Val Phe Leu Gln Ile Arg Leu
85 90 95
55
       Arg Gln Ser Phe Asn Asn Leu Ile Pro Ser Leu Pro Phe Arg Ile Asp
100 105 110
       Asn Thr Lys Lys Ile Thr Glu Met Lys Lys Thr Thr Leu Thr Gly Ser
115 120 125
60
       Ile Cys Ala Leu Leu Leu Phe Leu Gly Leu Ser Ala Asn Ala Gln Ser
130 135 140
       Lys Leu Lys Ile Lys Ser Ile Glu Ala Ala Thr Thr Phe Ser Ser Ala
145 150 150 160
65
       Thr Ala Gly Asn Gly Phe Gly Gly Asn Ile Phe Gly Met Asp Met Ser
165 170 175
       Ile Arg Met Arg Val His His Ser Ile Leu Pro Glu Gly Leu Asp Phe 180 185 190
       Ser Val Gly Ile His Glu Arg Arg Ala His Trp Glu Glu Ala Gly Ser
195 200 205
70
       Pro Lys Leu Net Tyr Thr Asn Val Pro Ser Ile Ile Gly Ile Val Glu
210 215 220
       Lys Val Ile Val Phe Glu Asp Ala Glu Asp Phe Phe Asp Lys Lys Ala
225 230 235 240
75
       Leu Gly Arg Phe Leu Ile Ser Leu Gly Ile Ser Tyr Thr Lys His Leu
```

```
Gly Ala Tyr Trp Gly Trp Thr Asn Asp Ala His Ile Leu Phe Ser Pro
260 265 270
        Ile Pro Lys Ser Lys Val His Tyr Asp Thr Tyr Thr Arg Ala Gly Ser 275 280 285
  5
        Asp Leu Val Leu Gln Ser Glu Asp Val Ala Thr Val Ser Ash Gly Phe 290 295 300
        Ser Pro Gly Ile Gly Leu Lys Ser Ser Ile Trp Trp Lys Met Pro Ile
305 310 315 320
 10
        Lys Ser Lys Tyr Asp Phe Arg Leu Gly Phe Ser Leu Gly Tyr Glu Tyr 325 \hspace{1.5cm} 330 \hspace{1.5cm} 335
        Leu Asn Leu Leu Tyr Pro Tyr Arg Asn Phe Lys Leu Asp Gly Asn Lys
340 345
        Pro Leu Ser Ala Leu Ser Pro Arg Met Asn His Ile Gly His Val Gly 355 360 365
 15
        Phe Asn Phe Thr Val Gly Leu Trp Thr Asn
        (2) INFORMATION FOR SEQ ID NO: 375
20
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1269 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
25
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
35
                   (B) LOCATION 1...1269
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:375
       40
       Phe Pro Phe Cys Ser Asp Tyr Tyr Thr Phe Glu Gly Val Thr Phe Leu 20 25 30
       Cys Ala Ser Asp Asp Met Thr Thr Lys Lys Pro Gln Ala Ile Leu Asp
35 40
45
       Leu Glu Lys Ala Tyr Asn Ile Glu Ile Pro Asp Leu Ser Ser Gln Glu
50 55 60
       Gly Ile Ser Trp Ser Val Asn Arg Tyr Phe Lys Gln Asp Ser Ser Gly
65 70 80
       Ala Val Val Glu Leu Cys Leu Arg Glu Cys Gln Ile Glu Ser Met Thr 85 \hspace{0.5cm} 90 \hspace{0.5cm} 95
50
       Trp Leu Ile Asp Phe Pro Ala Leu Lys Lys Leu Asp Leu Ser Tyr Asn 100 105 110
       Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr Lys
115 120 125
       Leu Arg Leu Arg Ser Asn Gln Ile Arg Lys Leu Glu Gly Leu Asp Ser
130 135 140
55
       Leu Thr Ser Leu Thr Lys Leu Ser Leu Ser Asp Asn Gln Ile Ser Lys
145 150 155 160
       Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Glu Leu Tyr Leu Leu
165 170 175
60
       Asp Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu
180 185 190
       Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu Gly Leu
195 200 205
65
       Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile
210 220
       Arg Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr Lys Leu Arg
225 230 235 240
       Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr 245 \hspace{1cm} 250 \hspace{1cm} 255 \hspace{1cm}
70
       Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu
260 265 270
       Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn
                                       280
75
       Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Ser Ser Leu Thr Lys
```

WO 99/29870 PCT/AU98/01023

| | | 290 |) | | | | 295 | • | | | | 300 |) | | | |
|------------|--------------|-----|-----|-----|------------|-----|-----|-----|-----|------------|------------|-----|-----|------------|------------|-----|
| | 305 | | | | | 310 |) | | | | Leu 315 | Glu | G13 | | | 320 |
| 5 | | | | | 325 |) | | | | 330 | Asp | | | | 335 | |
| | | | | 340 |) | | | | 345 | , | Ala | | | 350 | | |
| 10 | | | 355 |) | | | | 360 | | | Glu | | 365 | | | |
| 10 | | 370 |) | | | | 375 | | | | Ser | 380 | | | | |
| | 385 | | | | | 390 |) | | | | Leu 395 | | | | | 400 |
| 15 | | | | | 405 | | | | | 410 | Ser | | | | 415 | |
| | | | | 420 | | | | | 425 | | Gly | | | 430 | | |
| 20 | | | 435 | | | | | 440 | | | Gln | | 445 | | | |
| 20 | | 450 | | | | | 455 | | | | Leu | 460 | | | | |
| | 465 | | | | | 470 | | | | | Leu 475 | | | | | 480 |
| 25 | | | | | 485 | | | | | 490 | Asn | | | | 495 | |
| | | | | 500 | | | | | 505 | | His | | | 510 | | |
| 0.0 | | | 515 | | | | | 520 | | | Lys | | 525 | Ser | | |
| 30 | | 530 | | | | | 535 | | | | Glγ | 540 | | | | _ |
| | 545 | | | | | 550 | | | | | Asn 555 | | | | | 560 |
| 35 | | | | | 565 | | | | | 570 | Asn | | | | 575 | Ile |
| | | | | 580 | | | | | 585 | | His | | | 590 | | |
| 40 | | | 595 | | | | | 600 | | | Phe | | 605 | | | |
| 4 0 | | 610 | | | | | 615 | | | | Glu | 620 | | | | |
| | 625 | | | | | 630 | | | | | Ala 635 | | | | | 640 |
| 45 | | | | | 645 | | | | | 650 | Lys | | | | 655 | Thr |
| | | | | 660 | | | | | 665 | | Ala | | | 670 | Gly | |
| 5 0 | | | 6/5 | | | | | 680 | | | Asn | | 685 | | | |
| 50 | | 690 | | | | | 695 | | | | Ser | 700 | | | | |
| | 705 | | | | | 710 | | | | | Val 715 | | | | | 720 |
| 55 | | | | | /25 | | | | | 730 | Leu | | | | 735 | |
| | | | | 740 | | | | | 745 | | lle | | | Glu 750 | Ala | |
| co | Ala | | 155 | | | | | 760 | | | | | 765 | | | |
| 60 | | //0 | | | | | 775 | | | | Ser | 780 | | | | |
| | Leu 785 | | | | | 790 | | | | | 795 | | | | | 800 |
| 65 | | | | | 805 | | | | | 810 | Glu | | | | 815 | |
| | Asn | | | 820 | | | | | 825 | | | | | 830 | | |
| 70 | His . | | 835 | | | | | 840 | | | | | 845 | | | |
| 70 | | 850 | | | | | 855 | | | | | 860 | Val | | | |
| | Lys . 865 | | | | | 870 | | | | | Pro 875 | Leu | | | | 880 |
| 75 | Glu . | Ala | Tyr | Lys | Trp 885 | Leu | Thr | Leu | Gly | Phe 890 | Glu | Arg | Pro | | Phe 895 | Val |
| | | | | | | | | | | | | | | | 555 | |

```
Leu Lys Phe Glu Arg Phe Ile Pro Phe Gly Leu Ile As<br/>n Gl<br/>n Ile Ile 900 905 910
        Ala Tyr Tyr Gly Arg Glu Glu Gly Ala Leu Lys Arg Tyr Trp Arg Asp
915 920 925
 5
        Gln Val Ile Phe Thr Ala Gly Arg Glu Het Asp Arg Gln Thr Leu Glu
930 935 940
        Gln Glu Glu Glu Lys Glu Gly Leu Pro Lys Thr Asn Ala Glu Asp Tyr
945 950 955 960
        Gln Ile Trp Ile Lys Leu Asp Phe Thr Asp Leu Ala Ile Ser Val Phe
965 970 975
10
        Ile Lys Glu Gln Arg Lys Thr Ser Ala Lys Asp Met Gln Arg Lys Glu
980 985 990
        Ala Thr Ile Leu Ser Asp Met Leu Asp Met Tyr Trp Asn Asn Ile Pro
995 1000 1005
15
        Pro Arg Glu Gln Ile Gly Asp Lys Asp Thr Glu Gln Thr Arg Ser Thr 1010 1020
        Ile Arg Glu Thr Asn Arg Lys Lys Arg Pro Ile Gln Asp Leu Tyr Leu
1025 1030 1035
                                                          1035
        Ser Cys Ala Gln Ala Asp Lys Asp Leu Thr Glu Ser His Tyr Ile His
1045 1050 1055
20
        Leu Gly Thr Leu Asp Asp Glu Ser Lys Thr Thr Ala Arg Ile Ala Ala
1060 1065 1070
        Tyr Pro Leu Lys Asn Gly Val Ile Asp Lys Glu Arg Val Arg Glu Val
1075
1080
1085

Ser Thr Arg Fro Tyr Lys His Leu Ser Val Asn Lys Asn Leu Ala Thr
1090
1095
1100
25
        Ala Lys Gln Ile Phe Ile Ser Tyr Ser Lys Glu Asp Gln Thr Glu Leu
1105 1110 1115
        Glu Thr Cys Leu Gln Phe Phe Lys Pro Leu Glu Lys Asn Gly Gln Ile
1125 1130 1135
30
        Glu Ile Tyr Tyr Asp Lys Leu Thr Lys Phe Glu Thr Pro Ile His Pro
1140 1145 1150
        Glu Ile Arg Lys Arg Ile Val Glu Ala Asp Cys Ile Ile Ala Leu Ile
1155 1160 1165
        Ser Gln Arg Tyr Leu Ala Thr Asp Tyr Ile Leu Asp His Glu Leu Pro
1170 1175 1180
35
        Val Phe Arg Glu Tyr Asn Lys Thr Ile Val Pro Ile Leu Ile Lys Pro
1185 1190 1195
        Cys Thr Phe Glu Asp Asp Glu Phe Leu Arg Glu Lys Tyr Phe Ala Gln
1205 1210 1215
40
        Lys Ala Gln Ile Ile Asn Leu Gly Lys Glu Gly Lys Thr Ile Lys Ala
1220 1225 1230
        Tyr Asp Ser Ile Thr Ala Ser Ala His Arg Asp Glu Asn Trp Val Ala
1235 1240 1245
45
        Val Val Arg Glu Phe Lys Glu Lys Ile Leu Arg Ile Thr Lys Gln Glu
1250 1255 1260
                                    1255
                                                               1260
        Val Asn Thr Asp Glu
50
        (2) INFORMATION FOR SEQ ID NO: 376
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 231 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
60
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
65
                    (A) NAME/KEY: misc_feature (B) LOCATION 1...231
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:376
       Lys Phe Het Het Lys Lys Ala Phe Val Phe Val Leu Leu Val Cys Leu 1 5 15
70
       Phe Ser Ser Phe Ser Ser Ser Ala Gln Thr Thr Thr Asn Ser Ser Arg
20 25 30
       Ser Tyr Phe Thr Gly Arg Ile Glu Lys Val Ser Leu Asn Leu Gly Val
75
                                       40
```

```
Pro Pro Val Ser Thr Glu Val Trp Gly Met Thr His Asp Ala Asn Gly 50 55 60
        Leu Pro Phe Glu Ile Pro Ile Ser Phe Ser Arg Phe Asn Ser Gln Gly 65 70 75 80
       Asp Ile Ala Thr Thr Tyr Ile Ala Asn Ser Glu Ala Thr Leu Asn 85 90 95
       Glu Trp Cys Asp Tyr Ala His Pro Gly Gly Ile Val Arg Val Glu Gly
100 105 110

Arg Phe Trp Lys Met Thr Tyr Asn Ile Pro Thr Tyr Asn Ala Val Cys
115 120 125
10
       Thr Arg Ile Thr Phe Glu Asn Gln Glu Ile Glu Gly Thr Ile Val Leu
130 135 140
       Ile Pro Lys Pro Lys Val Ser Leu Pro His Val Ser Glu Ser Val Pro 145 155 160
15
       Cys Ile Arg Thr Glu Ala Gly Arg Glu Phe Ile Leu Cys Glu Glu Asp 165 170 175
       Asp Thr Phe Val Ser His Asp Gly Asn Glu Val Thr Ile Gly Gly Lys
180 185 190
       Pro Phe Leu Leu Asn Thr Asn Val Lys Ile Val Gly Asp Val Ser Gln 195 200 205
20
       Lys Tyr Ala Val Gly Val Gly Glu Ile Arg Phe Leu Gln Ile Cys Ala
210 215 220
       Gln Thr Val Ser Gln Gln Lys
25
        (2) INFORMATION FOR SEQ ID NO: 377
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 563 amino acids
30
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
35
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...563
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:377
45
       Ser Leu Ile Asn Asn Ser Arg Asp Trp Arg Ala Pro Val Arg Leu Ser 1 \phantom{\bigg|} 10 \phantom{\bigg|} 15
       Leu Lys Lys Thr Lys Thr Met Lys Thr Lys Val Leu Arg Lys Phe 20 25 30
50
       Val Val Ala Ala Phe Ala Val Ala Thr Leu Cys Pro Leu Ala Gln Ala 35 40 45
       Gin Thr Met Gly Gly Asp Asp Val Lys Val Val Gin Tyr Asn Gln Glu
50 55 60
       Lys Leu Val Gln Thr Arg Met Ser Val Ala Asp Asn Gly Trp Ile Tyr
65 75 80
55
       Val Met Thr His Ser Gly Tyr Asp Thr Gly Asn Ser Asn Val Lys Ile
85 90 95
       Phe Arg Ser Lys Asp Gln Gly Ala Thr Tyr Gln Lys Leu Arg Asp Trp
100 105 110
       Asp Pro Ser Asp Asp Tyr Gln Phe Gln Asp Phe Asp Ile Val Val Thr
115 120 125
60
       Gly Lys Asn Glu Ser Asp Ile Lys Ile Trp Ser Val Glu Leu Met Asn
130 135 140
       Lys Pro Gly Gly Tyr Lys Ser Arg Val Ala Val Phe Ser Arg Asp Ala
145 155 160
       Asn Ala Gln Asn Ala Lys Leu Val Tyr Lys Glu Asp Phe Ser Asn Val
165 170 175
       Gln Leu Tyr Asp Val Asp Ile Ala Ser Asn Tyr Arg Ser Pro Ser Ser
180 185 190
70
       Leu Asn Asn Gly Gly Asn Pro Phe Ala Leu Ala Phe Ala Tyr Thr Gly
195 200 205
       Phe Asn Asn Thr His Lys Ile Ser Phe Val Asp Tyr Val Phe Ser Leu
210 215 220
       Asn Gly Gly Gln Asn Phe Asn Lys Asn Leu Leu Phe Ser Gln Asp Gly
75
```

```
Glu Lys Lys Ile Asp Lys Val Asp Leu Ser Leu Gly Ser Thr Ser Glu 245 250 255
       Ser Met Gly His Asn Ala Trp Pro Leu Met Gly Val Val Phe Glu Met 260 265 270
                    260
                                      265
       Asn Lys Gln Gly Gly Lys Ser Asp Ile Gly Phe Leu Ser Asn Phe Val
275 280 285
 5
       Asp Asn Asp Pro Glu Phe Gln Trp Ser Gly Pro Ile Lys Val Ser Glu
290 295 300
       Ser Asp Met Ser Phe Ser Pro Lys Ile Gln Met Leu Asp Glu Asp 305 310 315 320
10
       Asn Asn Thr Ile Asn Gly Glu Ser Cys His Asn Phe Met Ile Thr Tyr
325 330 335
       Ser Asp Tyr Asp Ser Glu Tyr Ser Asp Trp Asp Ile Arg Tyr Val Tyr
340 345 350
       Pro Lys Lys Ser Phe Lys Tyr Glu Lys Gly Lys Thr Pro Thr Met Asp 355 360 365
15
       Asp Leu Val Glu Ala Phe Leu Thr Ala Ser Tyr Gln Ser Glu Thr Asn
370 375 380
       Ser Gly Leu Gly Tyr Asp Lys Asn Ala Asn His Tyr Leu Ile Thr Tyr
385 390 395 400
20
       Ala Lys Lys Glu Glu Asn Gly Thr Asn Thr Leu Lys Tyr Arg Trp Ala
405 410 415
       Asn Tyr Asp Lys Ile His Asn Lys Asp Leu Trp Ser Asp Thr Phe Thr 420 425 430
       Tyr Thr Ser Ser Ala Asn Ala Leu Tyr Thr Pro Gln Val Asp Ile Asn 435 440 445
25
       Pro Thr Lys Gly Leu Val Cys Trp Ser Trp Val Glu Tyr Leu Pro Gly 450 455 460
       Lys Arg Ile Val Trp Ser Asp Thr Gln Trp Thr His Ala Asn Gly Val
465 470 475 480
30
       Glu Asp Ile Val Met Gln Glu Gly Ser Met Lys Leu Tyr Pro Asn Pro
485 490 495
       Ala Gl<br/>n Glu Tyr Ala Val Ile Ser Leu Pro Thr Ala Ala As<br/>n Cys Lys 500 \phantom{000}505\phantom{000}510
35
       Ala Val Val Tyr Asp Met Gln Gly Arg Val Val Ala Glu Ala Ser Phe 515 520 525
       Ser Gly Asn Glu Tyr Arg Leu Asn Val Gln His Leu Ala Lys Gly Thr 530 535 540
       Tyr Ile Leu Lys Val Val Ser Asp Thr Glu Arg Phe Val Glu Lys Leu
40
       Ile Val Glu
       (2) INFORMATION FOR SEQ ID NO:378
45
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 786 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
50
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
55
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
60
                   (B) LOCATION 1...786
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:378
       Ile Val Tyr Leu Cys His Cys Met Asn His Arg Arg Ser Lys Thr Met
65
                                                10
       Leu Thr Ile Arg Asn Phe Leu Leu Phe Cys Cys Leu Ser Leu Ile Ala
20 25 30
       Phe Ala Ala Asp Ala Gln Ser Ser Val Ser Ser Gly Arg Leu Thr
35 40 45
70
       Glu Tyr Val Asn Pro Phe Ile Gly Thr Ala Asn Tyr Gly Thr Thr Asn 50 60
       Pro Gly Ala Val Leu Pro Asn Gly Leu Met Ser Val Thr Pro Phe Asn 65 70 80
       Val Ser Gly Ser Thr Glu Asn Arg Phe Asp Lys Asp Ser Arg Trp Trp
85 90 95
75
```

| | Ser | Ala | Pro | T;·r | Ser | Ala | Asp | Asn | Ser | Tyr | C?.e | Ile | GJ7. | Phe | Ser | His |
|----|------------|------------|------------|------------|------------|------------|--------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | | | | 100 | | | | Cys | 105 | | | | | 110 | | |
| 5 | | | 115 | | | | | 120 Asp | | | | | 125 | | | |
| Ü | | 130 | | | | | 135 | Arg | | | | 140 | | | | |
| | 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| 10 | | | | | 165 | | | Ala | | 170 | | | | | 175 | |
| | Leu | Thr | Glu | Phe 180 | Ala | Phe | Pro | Glu | Gly 185 | Glu | Gly | His | Ile | Leu 190 | Leu | Asn |
| | Leu | Gly | Gln 195 | Ala | Leu | Ser | Asn | Glu 200 | Ser | Gly | Ala | Ser | Val 205 | Arg | Phe | Leu |
| 15 | Asn | Asp 210 | Ser | Thr | Val | Val | Gl ₃ . 215 | Ser | Arg | Leu | Het | Gly 220 | Thr | Phe | Сув | Tyr |
| | Asn 225 | | Gln | Ala | Val | Phe 230 | | Gln | Туг | Phe | Val 235 | Leu | Gln | Val | Ser | Arg 240 |
| 20 | | Pro | Ile | Ser | | | Τyr | Trp | Lys | Lys 250 | | Pro | Pro | Met | Thr 255 | |
| 20 | Glu | Ala | Gln | | 245 Asp | Ser | Thr | Ala | | | тут | Lys | Gln | | | Glÿ |
| | Tyr | Lys | | 260 Glu | Met | Ser | Gly | Asp | 265 Asp | Ile | Gly | Val | | 270 Phe | Ser | Phe |
| 25 | Asn | - | 275 Asp | Gln | Gly | Glu | | 280 Ile | Tyr | Val | Arg | | 285 Ala | Val | Ser | Phe |
| | Val | 290 Ser | Glu | Ala | Asn | Ala | 295 Leu | Tyr | Asn | Leu | Glu | 300 Ala | Glu | Gln | Glu | Glu |
| | 305 | | | | | 310 | | Asn | | | 315 | | | | | 320 |
| 30 | | | | | 325 | | | Trp | | 330 | | | | | 335 | |
| | | | | 340 | | | | Glu | 345 | | | | | 350 | | |
| 35 | | | 355 | _ | | | - | 360 | _ | | | | 365 | | | |
| 33 | | 370 | | | | | 375 | Asn | | | | 380 | | | | |
| | 385 | | | | _ | 390 | _ | Lys | | _ | 395 | | | | | 400 |
| 40 | | | | | 405 | | | Asp | | 410 | | | | | 415 | |
| | Leu | Суѕ | Leu | Leu 420 | Tyr | Pro | Glu | Lys | Gln 425 | Leu | Asp | Met | Val | Arg 430 | Thr | Leu |
| | Ile | Asp | Met 435 | Tyr | Arg | Glu | Ser | Gly 440 | Trp | Leu | Pro | Arg | Trp 445 | Glu | Leu | Tyr |
| 45 | Gly | Gln 450 | Glu | Thr | Leu | Thr | Met 455 | Glu | Gly | Asp | Pro | Ser 460 | Leu | Ile | Val | Ile |
| | Asn 465 | | Thr | Trp | Gln | Arg 470 | | Leu | Arg | Ala | Phe 475 | | Thr | Ala | Thr | Ala 480 |
| 50 | | Glu | Ala | Met | Lys 485 | | Asn | Ala | Ser | Ser 490 | | Gly | Ala | Thr | His 495 | |
| 00 | Ile | Arg | Pro | | | Asp | Asp | Tyr | | | Leu | Gly | Phe | | | Leu |
| | Arg | Glu | | 500 Туг | Asp | Asn | Ser | Val | 505 Ser | His | Ala | Leu | | 510 Tyr | Tyr | Leu |
| 55 | Ala | | 515 Trp | Asn | Leu | Ser | | 520 Phe | Ala | His | Ala | | 525 Gly | His | Lys | Glu |
| | | 530 Ala | Ala | Leu | Phe | | 535 Lys | Arg | Ser | Leu | | 540 Tyr | Arg | His | Tyr | |
| | 545 Asn | Lys | Glu | Tyr | Gly | 550 Met | Leu | Cys | Pro | Leu | 555 Leu | Pro | Asp | Gly | Ser | 560 Phe |
| 60 | Leu | Thr | Pro | Phe | 565 Asp | Pro | Lys | Gln | Gly | 570 Glu | Asn | Phe | Glu | Pro | 575 Asn | Pro |
| | Gly | Phe | His | 580 Glu | Gly | Ser | Ala | Туг | 585 Asn | Tyr | Ala | Phe | Phe | 590 Val | Pro | His |
| 65 | | | 595 | | | | | 600 Leu | | _ | | | 605 | | | |
| | • | 610 | | _ | | | 615 | Asp | | • | - | 620 | - | | | |
| | 625 | | | | | 630 | | | | | 635 | | | | | 640 |
| 70 | | | - | | 645 | - | | Tyr | | 650 | | _ | | | 655 | |
| | | _ | | 660 | | - | | Thr | 665 | | | | | 670 | | |
| | Cys | Asn | Ala 675 | Pro | Asn | Gly | Leu | Pro 680 | Gly | Asn | Asp | Asp | Ala 685 | Gly | Thr | Иet |
| 75 | Ser | Ala | Trp | Leu | Val | Tyr | Ser | Met | Leu | Gly | Phe | Tyr | Pro | Asp | Cys | Pro |

```
695
       Gly Ser Pro Thr Tyr Thr Leu Thr Ser Pro Val Phe Pro Arg Val Arg
705 710 715 720
       Ile Arg Leu Asn Pro Gln Tyr Tyr Pro Gln Gly Glu Leu Ile Ile Thr 725 730 735
       Thr Asn Thr Glu Asn Gln Pro Thr Asp Ser Ile Tyr Ile His Thr Val
740 750
       Ser Leu Gly Asn Lys Thr Leu Pro His Gly Thr Arg His Ile Ser His
755 760 765
       Ala Asp Leu Val Arg Cys Gly His Leu Arg Tyr Glu Leu Ser Asn Arg
       Pro Arg
15
       (2) INFORMATION FOR SEQ ID NO:379
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 814 amino acids
                   (B) TYPE: amino acid
20
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...814
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:379
       Pro Arg Phe His Pro Gly Gly Thr His Ala Gln Leu Cys Arg Asn Gly
1 10 15
       Arg Arg Asn Gln Lys Ser Asn Gln Ser Ser Cys Phe Gly Gly Gly Thr
20 25 30
       Ala Pro Arg Phe Phe Ile Met Cys Lys Ile Arg Phe Ser Leu Leu Gln 35 40
40
       Ala Leu Val Val Cys Leu Leu Phe Thr Ser Phe Ser Leu Gln Ala Gln 50 60
       Glu Glu Gly Ile Trp Asn Thr Leu Leu Ala Ile His Lys Thr Glu Lys 65 70 75 80
45
       Ala Val Glu Thr Pro Lys Lys Val Phe Ala Val Ala Asn Gly Val Leu
85 90 95
       Tyr Ser Val Gly Lys Glu Ala Pro His Glu Ala Lys Ile Phe Asp Arg
100 105 110
       Ile Ser Gly Leu Ser Asp Thr Ser Val Ser Ser Ile Ala Tyr Ser Glu
115 120 125
50
       Gln Leu Lys Ser Leu Val Ile Tyr Tyr Ala Ser Gly Asn Ile Asp Ile
130 135 140
       Leu Asp Glu Ala Gly Arg Val Thr Asn Val Pro Ala Leu Lys Asp Asn 145 150 160
55
       Ile Asp Leu Ile Asp Lys Thr Leu Asn Arg Leu Leu Ile Val Gly Asn
165 170 175
       Arg Ala Tyr Leu Ala Gly Gly Phe Gly Leu Ser Val Leu Asp Val Ala
180 185 190
       Glu Ala Arg Ile Pro Ala Thr Tyr Ala Lys Gly Thr Lys Val Thr Asp
195 200 205
60
       Val Ala Lys Leu Asp Asn Asp Arg Leu Leu Het Leu Lys Glu Gly Gln 210 215 220
       Leu Phe Ile Gly Lys Glu Thr Asp Asn Leu Gln Asp Pro Ala Ala Trp
225 230 235 240
65
       Thr Ala Leu Ser Leu Asn Leu Pro Met Gly Ser Val Thr Gly Leu Gly 245 250 255
       Ile Val Gly Glu Asp Ile Cys Phe Leu Leu Ala Asp Gly Arg Val Tyr
260 265 270
       Val Ala Ala Asn Gln Ser Phe Glu Pro Glu Leu Leu Ser Ser Ser
275 280 285
       Ala Asp Ser Arg Leu Tyr Val Thr Asp Arg Gly Leu Phe Ile Cys Ala
290 295 300
       Glu Asn Arg Ile Tyr Phe Ile Glu Lys Gly Arg Lys Thr Thr Gln Phe
75
       Pro Ile Ala Asp Val Leu Gly Val Gly Ala Met Asn Glu Ser Asn Thr
```

| | | | | | 325 | | | | | 330 | | | | | 225 | |
|----|------------|------------|-------------|------------------------------|-------------|--------------|------------|------------|---|------------|------------|------------|------------|------------|------------|------------|
| | Ala | Т/.г | Ile | Ala 340 | | Gly | Glu | Glu | Gly 345 | | | Ser | Leu | Leu 350 | 335 Leu | Ala |
| 5 | Glu | G17. | Ser 355 | Thr | Ala | Glu | Ala | Met 360 | Pro | | Ala | Phe | Asp 365 | | Pro | Gly |
| | Asp | Asn 370 | Asp | Phe | ТУг | Glu | Het 375 | Arg | Phe | Ser | His | G17 380 | Arg | Leu | туг | Ala |
| | Ala 385 | Ser | Gly | Leu | Trp | Gly 390 | | Asn | Leu | Met | Gly 395 | His | Ala | Gly | Met | Val 400 |
| 10 | | | | Asp | 405 | | | | | 410 | | | | | 415 | |
| | | | | Leu 420 | | | | | 425 | | | | | 430 | | |
| 15 | | | 435 | Asn | | | | 440 | | | | | 445 | | | |
| | | 450 | | Phe | | | 455 | | | | | 460 | | | - | |
| | 465 | ASN | GIU | Thr | ATA | 470 | Ala | GLu | CAE | Asn | 475 | GIY | Asp | Ala | Arg | Val 480 |
| 20 | Lys | Ala | Ιlε | Ala | Phe 485 | Asp | Asn | Гλг | Gly | Asn 490 | Leu | Trp | Gly | Thr | Leu 495 | G1 y |
| | | | | Lys 500 | | | | | 505 | | | | | 510 | | • |
| 25 | His | Ser | Phe 515 | Ser | Tyr | Pro | Asp | Val 520 | Ala | Asn | Leu | Ala | Ser 525 | Phe | Gly | Asn |
| | Met | 11e 530 | | Leu | Pro | Asn | Gly 535 | | $\mathbf{L}_{\mathbf{J}}^{*}\mathbf{s}$ | Trp | Val | Asn 540 | | Leu | His | Arg |
| | Ser 545 | G17. | Gly | Ser | Thr | Arg 550 | | Gly | Val | Leu | 11e 555 | | Asn | Asp | Arg | Gly 560 |
| 30 | | | | Thr | 565 | | | | | 570 | | | | | 575 | |
| | | | | Leu 580 | | | | | 585 | | | | | 590 | | |
| 35 | Ala | Val | Asp 595 | His | Asn | Gly | Ser | Val 600 | Trp | liet | Gly | Ser | Asp 605 | Ile | Gly | Ile |
| | | 610 | | Туr | | | 615 | | | | | 620 | Thr | | | |
| 40 | 625 | | | Arg | | 630 | | | | | 635 | | | | | 640 |
| 40 | Leu | Asp | Lys | Val | Thr 645 | Val | Thr | Asp | Ile | Val 650 | Val | Asp | Lys | Leu | Asn 655 | His |
| | | | | Ala 660 | | | | | 665 | Leu | | | | 670 | Glu | |
| 45 | Cys | Ser | Ly:≲ 675 | Ile | Leu | Ala | Gln | Phe 680 | Thr | Val | Glu | Asn | Ser 685 | Pro | Leu | Leu |
| | | 690 | | Ile | | | 695 | Ala | | | | 700 | Asn | | | |
| | Tyr 705 | Ile | Gly | Thr | Ala | Asp 710 | Gly | Leu | Met | Thr | Phe | Gln | Thr | G1 y | Thr | |
| 50 | | G1; | Ser | Ala | Ser 725 | | Leu | Asp | Gly | Val 730 | | Val | Туг | Pro | Asn 735 | 720 Pro |
| | Leu | Arg | Pro | Glu 740 | | Pro | Asp | Gly | Val 745 | | Ile | Ala | Gly | Leu 750 | Gln | Ala |
| 55 | Gly | | Ser 755 | Val | Lys | Ile | | Asp 760 | | Thr | Gly | | Leu 765 | Leu | Tyr | Gln |
| | Thr | Glu 770 | Ser | Val | Thr | Thr | Glu 775 | Val | Lys | Trp | Asn | Ala 780 | Arg | Gly | Ala | Asp |
| | Gly 785 | Asn | Arg | Val | Ala | Ser 790 | Gly | Val | Туг | Ala | Val 795 | Ala | Val | Tyr | Asp | |
| 60 | | Ser | Lys | Lys | Ser 805 | | Leu | Ile | Arg | Phe 810 | | Val | lle | Arg | | 800 |
| | (2) | INFO | RHAT | NOI | FOR | SEQ | ID 11 | 10:38 | 10 | | | | | | | |
| 65 | | | | | | | | | | | | | | | | |
| 00 | | (1) | (A | UENC) LE) TY) TO | NGTH PE: | : 11 amin | 62 a | minc id | | .ds | | | | | | |
| 70 | | (ii) | HOL | ECUL | Е ТҮ | PE: | prot | ein | | | | | | | | |
| | (| iii) | НХЪ | отне | TICA | L: Y | ES | | | | | | | | | |
| 75 | | (vi) | | GINA) OR | | | | hyro | mona | s gi | ngiv | alis | | | | |

(ix) FEATURE:

(A) NAME/KEY: misc_feature (B) LOCATION 1...1162

| | | (xi |) SE | QUEN | CE DI | SCR | I PTI | эн: : | SEQ I | ED HO | 0:38 |) | | | | |
|------------|----------|-----|-------------------|------|-----------|-------|-------|-------|-------|-----------|------|-----|-----|-----|-----------|-----|
| | Ala 1 | Ile | Ser | Gln | Met 5 | Lys | Arg | Ile | Leu | Pro 10 | Ile | Val | Ala | Phe | Leu 15 | Ser |
| 10 | | | Leu | 20 | | | | | 25 | _ | | | | 30 | | _ |
| | _ | | Ala 35 | - | _ | | | 40 | | | - | | 45 | | • | |
| 15 | | 50 | Asp | | | | 55 | = | - | | | 60 | | | | _ |
| | 65 | | Gly | | | 70 | | | | | 75 | | | | | 80 |
| 00 | Glγ | Gly | Val | Leu | Pro 85 | Glu | Asp | Leu | Ser | Arg 90 | Ile | Thr | Thr | Asp | Asp 95 | Leu |
| 20 | | | Val | 100 | | | | | 105 | | | | - | 110 | | |
| | | | Pro 115 | | | | | 120 | | | | = | 125 | | | |
| 25 | | 130 | Val | | | | 135 | | | _ | _ | 140 | | | | _ |
| | 145 | | G1 y | | | 150 | | | | | 155 | | | _ | _ | 160 |
| 20 | | | Glu | | 165 | | | | | 170 | | | | | 175 | |
| 30 | | | Leu | 180 | | | | | 185 | _ | | | | 190 | _ | |
| | | | Ser 195 | | | | | 200 | | | | | 205 | | | |
| 35 | | 210 | Arg | | | | 215 | | | | | 220 | | | | |
| | 225 | | Arg | | | 230 | | | | | 235 | | | | | 240 |
| 40 | | | Leu | | 245 | | | | | 250 | | | | | 255 | |
| 40 | | | Ser | 260 | | | | | 265 | | | | | 270 | | |
| | | | Asn 275 | | | | | 280 | | | | | 285 | | | |
| 45 | | 290 | Gly | | | | 295 | | _ | | | 300 | | | | |
| | 305 | | Leu | | | 310 | | | | | 315 | | | | | 320 |
| 50 | | | Pro | | 325 | | | | | 330 | | | | | 335 | |
| 50 | | | Glu | 340 | | | | | 345 | | | | | 350 | | |
| | | | Val 355 | | | | | 360 | | | | | 365 | | | |
| 55 | | 370 | Pro | | | | 375 | | | | | 380 | | | - | |
| | 385 | | Leu | | | 390 | | | | | 395 | | _ | | | 400 |
| 60 | | | Asn | | 405 | | | | | 410 | | | | | 415 | |
| 00 | | | Ala | 420 | | | | | 425 | | | | | 430 | | |
| | | | 435 | | | | | 440 | | | | | 445 | | | |
| 65 | | 450 | Phe | | | | 455 | | | | | 460 | | | | |
| | 465 | | Met | | | 470 | | | | | 475 | | | | | 480 |
| 70 | | | Pro | | 485 | | | | | 490 | | | | | 495 | |
| , 3 | | | Val | 500 | | | | | 505 | | | | | 510 | | |
| | | | Leu 515 Asp | | | | | 520 | | | | | 525 | | | |
| <i>7</i> 5 | | 530 | , P | ייבט | - y - | F 114 | 535 | n-cu | บะน | veb | veft | 540 | £10 | U14 | 261 | AGI |



| | | Ile | Gly | Trp | Arg | Asn | Tyr | Asn | Met | Ala | Val | Gly | Arg | Phe | Pro | Val |
|------------|------------|------------|------------|------------|-------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | 545 Ara | Thr | Pro | Ala | Glu | 550 Ala | Ara | Ile | Ala | Val | 555 Asp | Lvs | Thr | Ile | Ara | 560 Tvr |
| = | | | | | 565 | | | | | 570 | | | | | 575 | |
| 5 | | | | 580 | | | | | 585 | Arg | | | | 590 | | |
| | Ala | Asp | Asn 595 | Gly | Asp | Lys | His | Ala 600 | Thr | Glu | Thr | Ser | Arg 605 | Leu | Ile | Asp |
| 10 | Thr | Val 610 | Lys | Arg | Tyr | Ala | Pro 615 | Ala | Ile | Met | Pro | Val 620 | Arg | Ala | Phe | Gln |
| | Asp 625 | Val | Tyr | Pro | His | Val 630 | Ile | Glu | Asn | Gly | Leu 635 | His | Ser | Ile | Pro | Gly 640 |
| | | Lys | Lys | Lys | Met 645 | | Glu | Thr | Leu | Gln 650 | | Gly | Ile | Ile | Leu 655 | |
| 15 | Asn | Tyr | Ala | Gly 660 | | Gly | Gly | Pro | Ala 665 | Gly | Trp | Ser | Asp | Glu 670 | | Leu |
| | Leu | Thr | Leu 675 | | Asp | Ile | His | Lys 680 | | Asn | Tyr | Lys | His 685 | | Pro | Ile |
| 20 | Trp | Ile 690 | | Ala | Thr | Cys | Asp 695 | | Ala | Asn | Tyr | Asp 700 | | Gln | Thr | Thr |
| 20 | | | Gly | Glu | Glu | | | Leu | His | Glu | | | Gly | Thr | Pro | |
| | 705 Het | Phe | Ser | Thr | Thr | 710 Arg | Val | Val | Tyr | Asn | 715 Thr | Gln | Asn | Glu | Lys | 720 Ile |
| 25 | | | | | 725 | | | | | 730 Glu | | | | | 735 | |
| | | | | 740 | | | | | 745 | Ser | | | | 750 | | |
| | | | 755 | | | | | 760 | | | | | 765 | | | |
| 30 | | 770 | | | | | 775 | | | Gln | | 780 | | | | |
| | 785 | | | | | 790 | | | | Pro | 795 | | | | | 800 |
| | Thr | Ala | Ile | Asn | Gly 805 | Gln | Asp | Pro | Glu | Gly 810 | Gln | Туг | Gly | Thr | Ile 815 | Met |
| 35 | Leu | Lys | Ser | Leu 820 | Glu | Arg | Val | Ala | Leu 825 | Lys | Gly | Lys | Val | Thr 830 | Asp | Glu |
| | Lys | Gly | Thr 835 | Phe | Asp | Glu | Thr | Phe 840 | Ser | Glу | Lys | Val | Phe 845 | Leu | Thr | Val |
| 40 | Phe | Asp 850 | | Arg | r7.2 | Lys | Met 855 | | Ala | Leu | Glu | Glu 860 | | Gly | Asn | Asp |
| | Leu 865 | Ser | Leu | Val | Tyr | Tyr 870 | | Tyr | Pro | Asn | Val 875 | | Tyr | Ala | Glу | Ile 880 |
| | Ala | Glu | Val | Lys | Asp 885 | Gly | Leu | Phe | Glu | Thr 890 | | Phe | Ile | Val | Pro 895 | |
| 4 5 | Asp | Val | Asn | Tyr 900 | | Glu | His | Glu | Gly 905 | Arg | Ile | Asn | Leu | Tyr 910 | | Tyr |
| | Asn | Glu | Ser 915 | | Lys | Ala | Glu | Ala 920 | | Gly | Val | Asp | Phe 925 | | Ile | Arg |
| 50 | Val | Gln 930 | | Gly | Ile | Pro | Asp 935 | | Val | Thr | Glu | | | Thr | Pro | Pro |
| 00 | | | Ile | Ser | Суѕ | | | Asn | Asp | Ser | | 940 Phe | Arg | Ser | Gly | |
| | 945 Glu | Val | Asn | Pro | | 950 Pro | Leu | Phe | Met | Ala | 955 Glu | Val | Phe | Asp | Leu | 960 Asn |
| 55 | Gly | Ile | Asn | Ile | 965 Thr | Gly | Ser | Gly | Val | 970 Gly | His | Asp | Ile | Thr | 975 Leu | Суѕ |
| | Ile | Asp | Gly | 980 Arg | Ala | Asp | Leu | Thr | 985 Tyr | Asn | Leu | Asn | Ala | 990 Tyr | Phe | Thr |
| | Ser | Ser | 995 Ala | Thr | Asp | Ala | Glγ | 1000 Val | | Thr | Ile | Leu | 100 Phe |)5 Met | Ile | Pro |
| 60 | | 1010 |) | | | | 101 | 15 | | Arg | | 10 | 20 | | | |
| | 1025 | 5 | | | | 103 | 30 | | | Ser | 10 | 35 | | | _ | 1040 |
| 65 | | | | | 1045 | 5 | | | | 105 Leu | 0 | | | | 10 |)55 |
| | | | | 1060 |) | | | | 106 | 55 | | | | 10 | 70 | |
| | | | 1075 | 5 | | | | 108 | 10 | Asn | | | 10 | 85 | | |
| 70 | | 1090 |) | | | | 109 | 95 | | Thr | | 11 | .00 | | | |
| | 1105 | 5 | | | | 111 | . 0 | | | Tyr | 11 | Glu 15 | Pro | | | 1120 |
| | Lys | Trp | Asp | Leu | Thr 1125 | Ser | | Tyr | Glу | Val 113 | Lys | Ile | Gly | Asn | | Phe 35 |
| 75 | Tyr | Leu | Tyr | Arg | | | Val | Asn | Ser | Pro | | Gly | Gln | Thr | | |

| | Met | Ala | Lys 115 | 114 Lys 5 | | Ile | Val | Val | | | | | | 1 | 150 | |
|------------|--------------------|------------|------------|-----------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 5 | (2) | INF | ORMA' | LION | FOR | SEQ | IDI | NO: 3 | 81 | | | | | | | |
| | | (ì |) SE | QUEN | CE CI | HARA | CTER: | ISTI | CS: | | | | | | | |
| | | | | A) LI 3) Ti | | | | | acio | ds. | | | | | | |
| 10 | | | (1 | D) T | OPOL | OGY: | lin | ear | | | | | | | | |
| | | (ii |) 140 | LECU1 | LE T | YPE: | pro | tein | | | | | | | | |
| 15 | | (iii |) HYI | РОТН | ETICA | AL: | YES | | | | | | | | | |
| | | (vi | | IGINA A) OI | | _ | | ohvr | omon: | ae u | ingi | vali. | = | | | |
| | | (iv | | ATURI | | | | , iiy L | J.11.011. | ab g. | riigi | vall. | - | | | |
| 20 | | (17 | (7 | A) 117 B) LO | A IE/I | | | | atur | 9 | | | | | | |
| | | (xi |) SE(| 2UEN(| CE DE | ESCR: | [PTI | : 14C | SEQ : | ID NO | 38: | 1 | | | | |
| 25 | Phe 1 | Tyr | Thr | Glγ | Ile | Asp | Leu | His | Ile | Glu 10 | Ser | Gln | Met | Lys | Lys 15 | Leu |
| | | Pro | Leu | Leu 20 | Fen | Leu | Ile | Leu | | | Leu | Vāl | Gly | | Gly | Lys |
| 30 | Lys | Glu | Lys 35 | | Ser | Val | Thr | Glu 40 | 25 Ile | Ala | Arg | Glu | Lys 45 | 30 Lys | Arg | Ile |
| 00 | Thr | | | Leu | Tyr | Glu | | | Leu | Pro | Thr | | | Val | Lys | Gln |
| | Leu 65 | 50 Tyr | Glu | Asn | Ser | | 55 Gln | Asn | Lys | Asn | | 60 Val | Gly | Gln | Met | |
| 35 | | Ala | Ile | Glu | | 70 Gly | Lys | Arg | Met | | 75 Asn | Иet | Ser | Gln | Tyr | 80 Thr |
| | Asp | Ala | Met | | 85 Tyr | His | Gln | Lys | | 90 Leu | Asn | Ala | Ala | | 95 Asn | Leu |
| 40 | Arg | Asp | | 100 Ile | Va1 | Ala | Ala | | 105 Ala | Trp | Asn | His | | 110 Gly | Thr | Asp |
| 40 | Ser | | 115 Arg | Ile | Gly | Ala | | 120 Ala | Glu | Ala | Ser | | 125 Tyr | His | Tyr | Lys |
| | | 130 Leu | Ser | Leu | Ile | | 135 Ser | Phe | Ser | Gly | | 140 Gln | Asn | Arg | Pro | |
| 4 5 | 145 Ile | Lys | Ala | Arg | | 150 Ala | Ala | Leu | Asn | | 155 Ile | Gly | Asn | Ile | Asn | 160 Leu |
| | Glu | Leu | Glу | | 165 His | Asp | Glu | Ala | | 170 Lys | Asn | Phe | Leu | | 175 Ala | Leu |
| EO | Gln | Gly | | 180 Lys | Glu | Leu | Asp | | 185 Pro | Leu | Gly | Gln | | 190 Ile | Asn | Tyr |
| 50 | Ala | | 195 Leu | Gly | Arg | Ile | | 200 Arg | Gln | Arg | Lys | Glu | 205 Tyr | Asp | Lys | Ala |
| | Arg | 210 Thr | Tyr | Phe | Leu | Leu | 215 Ser | Leu | Glu | Gln | Asn | 220 Asn | Met | Ala | Glu | Asn |
| 55 | 225 L eu | Met | Gly | Ile | Gly | 230 Leu | Cys | Ser | Ile | Asn | 235 Leu | Gly | Glu | Val | Asp | 240 Glu |
| | | | | | 245 | | | | | 250 | | | | | 255 Ala | |
| • | Lys | Leu | Met | 260 Glu | Gln | Leu | Ser | Asp | 265 Arg | Trp | His | Trp | Leu | 270 Asn | Ser | Суз |
| 60 | | | 275 | | | | | 280 | | | | | 285 | | Leu | |
| | | 290 | | | | | 295 | | | | | 300 | | | Asn | |
| 65 | 305 | | | | | 310 | | | | | 315 | | | | Leu | 320 |
| | | | | | 325 | | | | | 330 | | | | | 335 Ser | |
| | | | | 340 | | | | | 345 | | | | | 350 | lle | _ |
| 70 | | | 355 | | | | | 360 | | | | | 365 | | Leu | |
| | | 370 | | | | | 375 | | | | | 380 | | | Phe | |
| 75 | 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| 7.0 | r≃u | ₽en | 3=1 | INE | ren | rne | WIB | Leu | rue. | тте | ser | 11e | ren | reu | Ile | ser |

| | | | | | 40E | | | | | 410 | | | | | 116 | |
|-----------|-----|------------|------------|-------------------|------------|-------|------------|------------|------------|------------|-------|------------|------------|------------|------------|-----|
| | Val | Leu | Thr | Tyr 420 | 405 Ala | Туг | Arg | Gln | Gly 425 | 410 Lys | Lys | His | Asn | Lуε 430 | 415 Leu | Ile |
| 5 | rle | Glu | Thr 435 | Asp | Lys | Leu | Arg | Ser 440 | | Phe | Phe | Thr | Gly 445 | | Thr | His |
| | Glu | Phe 450 | Arg | Thr | Pro | Ile | Thr 455 | Val | Ile | Gln | Gly | Leu 460 | Asn | Glu | r?.e | Het |
| 4.0 | 465 | | | Pro | | 470 | | | | | 475 | | | | | 480 |
| 10 | lle | 11e | Asp | Arg | Gln 485 | Ser | Ser | His | Met | Leu 490 | Asn | Leu | Vāl | Asn | Gln 495 | Leu |
| | | | | 500 500 | | | | | 505 | | | | | 510 | | _ |
| 15 | | | 515 | lle | | | | 520 | | | | | 525 | | | |
| | | 530 | | Gln | | | 535 | | | | | 540 | | | | |
| 00 | 545 | | | Val | | 550 | | | | | 555 | | | | | 560 |
| 20 | | | | Leu | 565 | | | | | 570 | | | | | 575 | |
| | | | | Ser 580 | | | | | 585 | | | | | 590 | | |
| 25 | | | 595 | Asp | | | | 600 | | | | | 605 | | | |
| | | 610 | | Ile | | | 615 | | | | | 620 | | | | |
| 20 | 625 | | | G1?. | | 630 | | | | | 635 | | | | | 640 |
| 30 | | | | L∵s | 645 | | | | | 650 | | | | | 655 | |
| | | | | Pro 660 | | | | | 665 | | | | | 670 | | |
| 35 | | | 675 | Ser | | | | 680 | | | | | 685 | | | |
| | | 690 | | Fro | | | 695 | | | | | 700 | | | | |
| 40 | 705 | | | Arg | | 710 | | | | | 715 | | | | | 720 |
| 40 | | | | Val | 725 | | | | | 730 | | | | | 735 | |
| | | | | Gly 740 | | | | | 745 | | | | | 750 | | |
| 45 | | | 755 | lle | | | | 760 | | | | | 765 | _ | | |
| | | 770 | | Arg | | | 775 | | | | | 780 | | | | |
| 50 | 785 | | | Thr | | 790 | | | | | 795 | | | | | 800 |
| 30 | | | | Val | 805 | | | | | 810 | | | | | 815 | |
| | | | | Arg 820 | | | | | 825 | | | | | 830 | | |
| 55 | | | 835 | Het | | | | 840 | | | | | 845 | | | |
| | | 850 | | Ile | | | 855 | | | | | 860 | | | | |
| 60 | 865 | | | Ser | | 870 | | | | | 875 | | | | | 880 |
| 00 | | | | Glu Wal | 885 | | | | | 890 | | | | | 895 | |
| | | | | Val 900 Leu | | | | | 905 | | | | | 910 | | |
| 65 | | | 915 | Ser | | | | 920 | | | | | 925 | | | |
| | | 930 | | r?.e | | | 935 | | | | | 940 | | | | |
| 70 | 945 | | | Ala | | 950 | | | | | 955 | | | 1 7 1 | Arg | 960 |
| - 🕶 | ٠,٠ | J#4 | 224 | | 965 | | | Ser | nep | 970 | o i u | 1111 | val | | | |
| | (2) | INFC | Ri IA1 | HOI | FOR | SEQ | ID N | 0:38 | 2 | | | | | | | |
| 75 | | (i) | SEÇ | DENC | E CH | IARAC | TERI | STIC | S: | | | | | | | |

```
(A) LENGTH: 563 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
 5
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
10
                    (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...563
15
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:382
       His His Lys Thr Tyr Gln Thr Met Lys Lys Leu Leu Gln Ala Lys Ala 1 5 10 15
20
       Leu Ile 1.eu Ala Leu Gly Leu Phe Gin Leu Pro Ala Ile Ala Gln Thr
20 25 30
       Gln Met Gln Ala Asp Arg Thr Asn Gly Gln Phe Ala Thr Glu Glu Met 35 40 45
       Gln Arg Ala Phe Gln Glu Thr Asn Pro Pro Ala Gly Pro Val Arg Ala
50 55 60
25
       Ile Ala Glu Tyr Glu Arg Ser Ala Ala Val Leu Val Arg Tyr Pro Phe 65 70 75 80
       Gly Ile Pro Met Glu Leu Ile Lys Glu Leu Ala Lys As<br/>n Asp Lys Val85 \hspace{1cm} 90 \hspace{1cm} 95
       Ile Thr Ile Val Ala Ser Glu Ser Gln Lys Asn Thr Val Ile Thr Gln
100 105 110
30
       Tyr Thr Gln Ser Gly Val Asn Leu Ser Asn Cys Asp Phe Ile Ile Ala
115 · 120 125
       Lys Thr Asp Ser Tyr Trp Thr Arg Asp Tyr Thr Gly Trp Phe Ala Net
130 135
35
       Tyr Asp Thr Asn Lys Val Gly Leu Val Asp Phe Ile Tyr Asn Arg Pro
145 150 160
       Arg Pro Asn Asp Asp Glu Phe Pro Lys Tyr Glu Ala Gln Tyr Leu Gly
165 170 175
       Ile Glu Het Phe Gly Met Lys Leu Lys Gln Thr Gly Gly Asn Tyr Het
180 185 190

Thr Asp Gly Tyr Gly Ser Ala Val Gln Ser His Ile Ala Tyr Thr Glu
195 200 205
40
       Asn Ser Ser Leu Ser Gln Ala Gln Val Asn Gln Lys Met Lys Asp Tyr 210 215 220
45
       Leu Gly Ile Thr His His Asp Val Val Gln Asp Pro Asn Gly Glu Tyr 225 230 235 240
       Ile Asn His Val Asp Cys Trp Gly Lys Tyr Leu Ala Pro Asn Lys Ile
245 250 255
50
       Leu Ile Arg Lys Val Pro Asp Asn His Pro Gln His Gln Ala Leu Glu 260 265 270
       Asp Het Ala Ala Tyr Phe Ala Ala Gin Thr Cys Ala Trp Gly Thr Lys
275 280 285
       Tyr Glu Val Tyr Arg Ala Leu Aia Thr Asn Glu Gln Pro Tyr Thr Asn 290 295 300
55
       Ser Leu Ile Leu Asn Arg Val Phe Val Pro Val Asn Gly Pro Ala
305 310 315 320
       Ser Val Asp Asn Asp Ala Leu Asn Val Tyr Lys Thr Ala Net Pro Gly
325 330 335
60
       Tyr Glu Ile Ile Gly Val Lys Gly Ala Ser Gl; Thr Pro Trp Leu Gly 340 345 350
       Thr Asp Ala Leu His Cys Arg Thr His Glu Val Ala Asp Lys Gly Tyr 355 360 365
       Leu Tyr Ile Lys His Tyr Pro Ile Leu Gly Glu Gln Ala Gly Pro Asp
370 375 380
65
       Tyr Lys Ile Glu Ala Asp Val Val Ser Cys Ala Asp Ala Thr Ile Ser 385 390 400
       Pro Val Gin Cys Tyr Tyr Arg Ile Asn Gly Ser Gly Ser Phe Lys Ala
405 415
70
       Ala Asp Het Thr Het Glu Ser Thr Gly His Tyr Thr Tyr Ser Phe Thr 420 425 430
       Gly Leu Ash Lys Ash Asp Lys Val Glu Tyr Tyr Ile Ser Ala Ala Asp 435 \  \  \, 440 \  \  \, 445
       Asn Ser Gly Arg Lys Glu Thr Tyr Pro Fhe Ile Gly Glu Pro Asp Pro
75
```

```
Phe Lys Phe Thr Cys Het Asn Glu Thr Asn Thr Cys Thr Val Thr Gly
465 470 475 480
       Ala Ala Lys Ala Leu Arg Ala Trp Phe Asn Ala Gly Arg Ser Glu Leu
                        485
                                           490
                                                                    495
       Ala Val Ser Val Ser Leu Asn Ile Ala Gly Thr Tyr Arg Ile Lys Leu 500 505 510
                   500
                                                        510
                                       505
       Tyr Asn Thr Ala Gly Glu Glu Val Ala Ala Het Thr Lys Glu Leu Val
515 520 525
       Ala Gly Thr Ser Val Phe Ser Met Asp Val Tyr Ser Gln Ala Pro Gly 530 535 540
10
       Thr Tyr Val Leu Val Val Glu Gly Asn Gly Ile Arg Glu Thr Met Lys
       Ile Leu Lys
15
       (2) INFORMATION FOR SEQ ID NO:383
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 437 amino acids
20
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
25
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Perphyromonas gingivalis
30
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...437
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:383
35
       Thr Thr Asn Arg Lys Pro Asn Thr Asn Met Lys Leu Ser Ser Lys Lys
                                              10
       Ile Leu Ala Ile Ile Ala Leu Leu Thr Met Gly His Ala Val Gln Ala 20 25 30
                                       25
40
       Gln Phe Val Pro Ala Pro Thr Thr Gly Ile Arg Met Ser Val Thr Thr
       35 40 45
Thr Lys Ala Val Gly Glu Lys Ile Glu Leu Leu Val His Ser Ile Glu
50 55 60
              35
       Lys Lys Gly Ile Trp Ile Asp Leu Asn Gly Asp Ala Thr Tyr Gln Gln 65 70 75 80
45
       Gly Glu Glu Ile Thr Val Phe Asp Glu Ala Tyr His Glu Tyr Thr Ile 85 90 95
       Gly Thr Gln Thr Leu Thr Ile Tyr Gly Asn Thr Thr Arg Leu Gly Cys
100 105 110
                                        105
50
       Arg Ser Thr Gly Ala Thr Ala Val Asp Val Thr Lys Asn Pro Asn Leu
115 120 125
       Thr Tyr Leu Ala Cys Pro Lys Asn Asn Leu Lys Ser Leu Asp Leu Thr 130 135 140
       Gln Asn Pro Lys Leu Leu Arg Val Trp Cys Asp Ser Asn Glu Ile Glu
145 150 155 160
55
       Ser Leu Asp Leu Ser Gly Asn Pro Ala Leu Ile Ile Leu Gly Cys Asp
165 170 175
       Arg Asn Lys Leu Thr Glu Leu Lys Thr Asp Asn Asn Pro Lys Leu Ala
180 185 190
60
       Ser Leu Trp Cys Ser Asp Asn Asn Leu Thr Glu Leu Glu Leu Ser Ala
195 200 205
       Asn Pro Arg Leu Asn Asp Leu Trp Cys Phe Gly Asn Arg Ile Thr Lys
210 215 220
       Leu Asp Leu Ser Ala Asn Pro Leu Leu Val Thr Leu Trp Cys Ser Asp 225 230 235 240
65
       Asn Glu Leu Ser Thr Leu Asp Leu Ser Lys Asn Ser Asp Val Ala Tyr
245 250 255
       Leu Trp Cys Ser Ser Asn Lys Leu Thr Ser Leu Asn Leu Ser Gly Val
260 265 270
70
       Lys Gly Leu Ser Val Leu Val Cys His Ser Asn Gln Ile Ala Gly Glu
275 280 285
       Glu Met Thr Lys Val Val Asn Ala Leu Pro Thr Leu Ser Pro Gly Ala
         290
                               295
                                           300
       Gly Ala Gln Ser Lys Phe Val Val Val Asp Leu Lys Asp Thr Asp Glu
75
```

```
Lys Asn Ile Cys Thr Val Lys Asp Val Glu Lys Ala Lys Ser Lys Asn 325 \hspace{1.5cm} 330 \hspace{1.5cm} 335
       Trp Arg Val Phe Asp Phe Asn Gly Asp Ser Asp Asn Met Leu Pro Tyr 340 345 350
       Glu Gly Ser Pro Thr Ser Asn Leu Ala Val Asp Ala Pro Thr Val Arg
355 360 365
       Ile Tyr Pro Asn Pro Val Gly Arg Tyr Ala Leu Val Glu Ile Pro Glu
370 375 380
       Ser Leu Leu Gly Gln Glu Ala Ala Leu Tyr Asp Met Asn Gly Val Lys
385 390 395 400
10
       Val Tyr Ser Phe Ala Val Glu Ser Leu Arg Gln Asn Ile Asp Leu Thr
405 415
       His Leu Pro Asp Gly Thr Tyr Phe Phe Arg Leu Asp Asn Tyr Thr Thr 420 425 430

Lys Leu He Lys Glo
15
       (2) INFORMATION FOR SEQ ID NO:384
20
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 318 amino acids
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
           (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
30
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...318
35
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:384
       Asn His Leu Lys Thr Asn Ile Lys Met Arg Lys Thr Ile Ile Phe Cys
                                                 10
       Leu Leu Leu Ala Leu Phe Gly Cys Ser Trp Ala Gln Glu Arg Val Asp
20 25 30
40
       Glu Lys Val Phe Ser Ala Gly Thr Ser Ile Phe Arg Gly Ile Leu Glu
35 40
       Lys Val Lys Ala Pro Leu Met Tyr Gly Asp Arg Glu Val Trp Gly Met 50 60
45
       Ala Arg Ala Ser Glu Asp Phe Phe Phe Ile Leu Pro Val Thr Asp Asp 65 70 80
       Leu Thr Pro Val Leu Phe Tyr Asn Arg Leu Thr Asn Glu Pro Cys Phe
85 90 95
50
       Val Ser Asp Gln Gly Ile Thr Glu Tyr Phe Lys Phe Ala Gln Glu Gl;
100 105
       Asp Tyr Ile Glu Val Glu Gly Ser Ser Val Phe Met Ala Asn Leu Leu
115 120 125
       Tyr Tyr Arg Phe Phe Pro Thr Arg Ile Thr Ser Tyr Asn Ala Pro Ile
130 135 140
55
       Glu Gly Val Val Ser Lys Thr Gly Asn Pro Ala Phe Thr Ile Pro Net 145 150 155 160
       Leu Pro Gly Val Ser Asp Cys Ile Glu Ile Ser Asn Asn Arg Lys Val
165 170 175
60
       Phe Leu Thr Asn Gln Leu Gly Val Val Asn Ile Thr Asp Gly Het Glu
180 185 190
       Pro Pro Ile Ile Ala Gly Val Ser Ala Ser Tyr Gly Ser Ser Val Arg
195 200 205
       Val Tyr Gly His Val Ser Gln Arg Trp Asp Ile Ile Gly His Cys Tyr
210 215 220
65
       Leu Asp Ile Tyr Pro Thr Asn Cys Tyr Pro Leu Ser Thr Lys Pro Val
225 230 235 240
       Ala Gly Asp Asp Glu Val Phe Val Lys Gln Gln Gly Arg Gln Ile Glu 245 250 255
70
       Ile Asp Ser Asn Ser Pro Ile Val Gln Val Val Tyr Asp Leu Glu
260 265 270
       Gly Lys Ser Val Phe Arg Lys Arg Het Thr Glu Asn Ala Tyr Thr Leu
275 280 285
       Ser Phe Arg Ala Pro Met Leu Gly Phe Met Thr Ile Met Ile Glu Thr
75
                                   295
```

```
Gln Asn Ser Ile Ile Asn Lys Lys Leu Asn Val Thr Gln Leu
                              310
        (2) INFORMATION FOR SEQ ID NO:385
  5
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 461 amino acids
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
 10
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
 15
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
20
                  (B) LOCATION 1...461
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:385
       Arg Arg Ala Val Lys lle Arg Ser Pro Pro His Ile His Ser Leu Phe
25
                                              10
       Val Arg Lys Cys Leu Phe Ser Asp Phe Lys Tyr Leu Tyr Leu Ser Arg 20 25 30
       Lys Ile Thr Glu Arg Leu Gly Arg Leu Ser Ile Arg Leu Lys Ser 35 40 45
30
       Tyr Asn Pro Ile Ile Ile Glu Met Lys Lys Thr Thr Ile Ile Ser
50 55 60
       Leu Ile Val Phe Gly Ala Phe Phe Ala Ala Val Gly Gln Thr Lys Asp
65 70 75
       Asn Ser Ser Tyr Lys Pro Phe Ser Lys Glu Asp Ile Ala Gly Gly Val
35
       Tyr Ser Leu Pro Thr Gln Asn Arg Ala Gln Lys Asp Asn Ala Glu Trp
100 105 110
       Leu Leu Thr Ala Thr Val Ser Thr Asn Gln Ser Ala Asp Thr His Phe
       Ile Phe Asp Glu Asn Asn Arg Tyr Ile Ala Arg Asp Ile Lys Ala Asn
130 135 140
40
       Gly Val Arg Lys Ser Thr Asp Ser Ile Tyr Tyr Asp Ala Asn Gly Arg
145 150 155
       Ile Ser His Val Asp Leu Tyr Ile Ser Phe Ser Gly Gly Glu Pro Ala
165 170 175
       Leu Asp Thr Arg Fhe Lys Tyr Thr Tyr Asp Asp Glu Gly Lys Het Thr
180 185
      Val Arg Glu Val Phe Met Leu Val Het Asp Pro Asn Thr Pro Ile Ser
195 200 205
Arg Leu Glu Tyr His Tyr Asp Ala Gln Gly Arg Leu Thr His Trp Ile
210 225
50
      55
      Thr Tyr Ser Asp Thr Gly Lys Thr Glu Tyr Ser Tyr Asp Asp Ala Asp 260 265 270
      Asn Het Val Lys Ala Glu Tyr Phe Val Val Gln Gln Gly Lys Ala Trp
275 280 285
60
      Gln Val Leu Lys Arg Glu Glu Tyr Thr Tyr Glu Asp Asn Ile Cys Ile
290 295 300
      Gln Tyr Leu Ala Ile Asn Gly Thr Asp Thr Lys Val Tyr Lys Arg Asp 305 310 315 320
      Ile Glu Ser Asp Lys Ser Ile Ser Ala Asn Val Ile Asp Ile Pro Ser
325 330 335
65
      Met Pro Glu Gln Thr Trp Pro Asn Met Tyr Gly Phe Asn Ala Lys Arg
340 345 350
      Leu Lys Glu Thr Tyr Ser Ser Tyr Glu Gly Asp Val Ala Thr Pro Ile
355 360 365
70
      Phe Asp Tyr Ile Tyr Thr Tyr Lys Ala Leu Thr Ser Het Ala Thr Pro
370 375
      Ser Thr Glu Ala Gln Val Ala Val Tyr Leu Asn Pro Ser Thr Asp Arg
385 390 395
      Leu Val Ile Leu Ala Asn Gly Ile Thr His Leu Ser Met Tyr Asp Leu
75
                                              410
```

```
Gln Gly Lys Leu 11e Arg Asp Cys Ala Leu Ser Gly Asp Lys Val Glu 420 425 430
       Het Gly Val Gly Ser Leu Thr Lys Gly Thr Tyr Leu Leu Lys Val Asn 435 440 445
 5
       Thr Asp Gln Gly Ala Phe Val Arg Lys Val Val Ile Arg
       (2) INFORMATION FOR SEQ ID NO:386
10
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 451 amino acids
                   (P) TYPE: amino acid
                   (D) TOPOLOGY: linear
15
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
20
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KET: misc feature
                   (B) LOCATION 1...451
25
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:386
       Het Ala Lys Val Ile Lys Thr Lys Lys Gly Leu Ala Leu Asn Leu Lys
                                                10
30
       Gly Lys Pro Leu Pro Glu Met Leu Ala Glu Pro Ala Gln Ser Pro Thr 20 25 30
       Tyr Ala Val Val Pro Asp Asp Phe Glu Gly Val Ile Pro Lys Val Thr 35 40
       Ala Arg Fro Gly Asp Lys Val Arg Ala Gly Ser Ala Leu Met His His 50 55 60
35
       Lys Ala Tyr Pro Glu Met Lys Phe Thr Ser Pro Val Ser Gly Glu Val 65 70 75 80
       Ile Ala Val Asn Arg Gly Ala Lys Arg Lys Val Leu Ser Ile Glu Val
40
       Lys Pro Asp Gly Leu Asn Glu Tyr Glu Ser Phe Pro Val Gly Asp Pro 100 \\ 0.05 \\ 100 \\ 110
       Ser Ala Leu Ser Ala Glu Gln Ile Lys Glu Leu Leu Leu Ser Ser Gly 115 120 125
       Het Trp Gly Phe Ile Lys Gln Arg Pro Tyr Asp Ile Val Ala Thr Pro
130 135 140
45
        130
       Asp Ile Ala Pro Arg Asp Ile Tyr Ile Thr Ala Asn Phe Thr Ala Pro 145 150 160
       Leu Ala Pro Asp Phe Asp Phe Ile Val Arg Gl\gamma Glu Glu Arg Ala Leu 165 170 175
50
       Gln Thr Ala Ile Asp Ala Leu Ala Lys Leu Thr Thr Gly Lys Val Tyr
180 185 190
       Val Gly Leu Lys Pro Gly Ser Ser Leu Gly Leu His Asn Ala Glu Ile
195 200 205
       Val Glu Val His Gly Pro His Pro Ala Gly Asn Val Gly Val Leu Ile
210 215 220
55
       Asn His Thr Lys Pro Ile Asn Arg Gly Glu Thr Val Trp Thr Leu Lys 225 230 235 240
       Ala Thr Asp Leu Ile Val Ile Gly Arg Phe Leu Leu Thr Gly Lys Ala
245 250 255
60
       Asp Phe Thr Arg Met Ile Ala Het Thr Gly Ser Asp Ala Ala Ala His
260 265 270
       Gly Tyr Val Arg Ile Het Pro Gly Cys Asn Val Phe Ala Ser Phe Pro
275 280 285
       Gly Arg Leu Thr Ile Lys Glu Ser His Glu Arg Val Ile Asp Gly Asn
290 295 300
65
       Val Leu Thr Gly Lys Lys Leu Cys Glu Lys Glu Pro Phe Leu Ser Ala
305 310 315 320
       Arg Cys Asp Gln Ile Thr Val Ile Pro Glu Gly Asp Asp Val Asp Glu
325 330 335
70
       Leu Phe Gly Trp Ala Ala Pro Arg Leu Asp Gln Tyr Ser Net Ser Arg 340 345 350
       Ala Tyr Phe Ser Trp Leu Gln Gly Lys Asn Lys Glu Tyr Val Leu Asp
355 360 365
       Ala Arg Ile Lys Gly Gly Glu Arg Ala Met Ile Met Ser Asn Glu Tyr
75
```

```
Asp Arg Val Phe Pro Met Asp Ile Tyr Pro Glu Tyr Leu Leu Lys Ala
                                                395
       Ile Ile Ala Phe Asp Ile Asp Lys Met Glu Asp Leu Gly Ile Tyr Glu
                       405
                                           410
                                                                415
 5
       Val Ala Pro Glu Asp Phe Ala Thr Cys Glu Phe Val Asp Thr Ser Lys
                                      425
                  420
                                                            430
       Ile Glu Leu Gln Arg Ile Val Arg Glu Gly Leu Asp Met Leu Tyr Lys
              435
                                    440
      Glu Met Asn
10
          450
       (2) INFORMATION FOR SEQ ID NO:387
            (i) SEQUENCE CHARACTERISTICS:
15
                 (A) LENGTH: 195 amino acids
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
20
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Perphyromonas gingivalis
25
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...195
30
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 387
      Het Asp Lys Val Ser Tyr Ala Leu Gly Leu Ser Ile Gly Asn Asn Phe
                                           1.0
      Lys Ser Ser Gly Ile Asp Ser Val Val Met Asp Asp Phe Met Gln Gly
35
                                       25
      Leu Ser Asp Val Leu Glu Glu Lys Ala Pro Gln Leu Ser Tyr Asp Glu
35 40 45
      Ala Lys Arg Glu Ile Glu Ala Tyr Phe Het Asp Leu Gln Gln Lys Ala
40
      Val Lys Leu Asn Lys Glu Ala Gly Glu Glu Phe Leu Lys Ile Asn Ala
65 70 75 80
      His Lys Glu Gly Val Thr Thr Leu Pro Ser Gly Leu Gln Tyr Glu Val
                       85
                                           90
      Ile Lys Het Gly Glu Gly Pro Lys Pro Thr Leu Ser Asp Thr Val Thr 100 105 110
45
      Cys His Tyr His Gly Thr Leu Ile Asn Gly Ile Val Phe Asp Ser Ser
115 120 125
      Het Asp Arg Gly Glu Pro Ala Ser Phe Pro Leu Arg Gly Val Ile Ala
          130
                             135
                                                  140
50
      Gly Trp Thr Glu Ile Leu Gln Leu Met Pro Val Gly Ser Lys Trp Lys
                          150
      Val Thr Ile Pro Ser Asp Leu Ala Tyr Gly Asp Arg Gly Ala Gly Glu
                      165
                                          170
                                                              175
      His Ile Lys Pro Gly Ser Thr Leu Ile Phe Ile Ile Glu Leu Leu Ser
55
      Ile Asn Lys
      (2) INFORMATION FOR SEQ ID NO:388
60
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 273 amino acids (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
65
          (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
70
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: Porphyromonas gingivalis
          (ix) FEATURE:
                (A) NAME/KEY: misc_feature
75
                 (B) LOCATION 1...273
```

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:388
       Met Lys Lys Ala Leu Leu Ile Gly Ala Ala Leu Leu Gly Ala Val Ser
 5
       Phe Ala Ser Ala Gln Ser Leu Ser Thr Ile Lys Val Gln Asn Asn Ser 20 25 30
       Val Gln Gln Pro Arg Glu Glu Ala Thr Ile Gln Val Cys Gly Glu Leu 35 40 45
10
       Ala Glu Gln Val Asp Cys Ile Gly Thr Gly Asn Ser Ala Ile Ile Ala 50 60
       Ala Ala Lys Phe Glu Ser Asp Asp Leu Glu Ser Tyr Val Gly Trp
65 70 75 80
       Glu Ile Net Ser Val Asp Phe Phe Pro Gly Tyr Lys Ala Cys Lys Tyr 85 90 95
15
       Thr Ser Ala Val Trp Ala Asp Asp Net Thr Ile Leu Gly Gln Ser Glu 100 105 110
       Asp Ser Asp Pro Glu Met Gln Thr Ile Asn Asn Leu Ala Leu Lys Thr 115 120 125
20
       Ser Val Lys Ile Glu Ala Gly Lys Asn Tyr Ile Val Gly Tyr Ile Ala
130 140
       Asn Thr Ala Gly Gly His Pro Ile Gly Cys Asp Gln Gly Pro Ala Val
145 150 160
       Asp Gly Tyr Gly Asp Leu Val Ser Ile Ser Glu Asp Gly Gly Ala Thr
165 170 175
25
       Phe Pro Pro Phe Glu Ser Leu His Gln Ala Val Pro Thr Leu Asn Tyr
180 185 190
       Asn Ile Tyr Val Val Val His Leu Lys Lys Gly Glu Gly Val Glu Ala 195 200 205
30
       Val Leu Thr Asn Asp Lys Ala Asn Ala Tyr Val Gln Asn Gly Val Ile
210 215 220
       Tyr Val Ala Gly Ala Asn Gly Arg Gln Val Ser Leu Phe Asp Met Asn 225 230 235 240
       Gly Lys Val Val Tyr Thr Gly Val Ser Glu Thr Ile Ala Ala Pro Gln 245 255 255

Lys Gly Met Tyr Ile Leu Arg Val Gly Ala Lys Ser Ile Lys Leu Ala 260 265 270
35
40
       (2) INFORMATION FOR SEQ ID NO:389
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 554 amino acids
45
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) HOLECULE TYPE: protein
50
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
55
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...5\overline{5}4
            (Mi) SEQUENCE DESCRIPTION: SEQ ID NO:389
60
       Het Pro Arg Ile Het Lys Leu Lys Ile Ala Leu Arg Leu Leu Leu Ala 1 5 10 15
       Thr Phe Ala Ile Val Leu Phe Ser Pro Leu Ala Lys Ala Gln Net Asp
20 25 30
       Ile Gly Gly Asp Asp Val Leu Ile Glu Thr Het Ser Thr Leu Ser Gly 35 40 45
65
       Tyr Ser Glu Asp Phe Tyr Tyr Lys Met Ala Val Ala Asp Asn Gly Trp 50 60
       Ile Tyr Val Met Leu Asp Phe Ser Arg Ile Tyr Phe Asp Asp Val Arg 65 70 75
70
       Leu Tyr Arg Ser Lys Asp Gly Gly Ala Thr Tyr Gln Lys Leu Gly Ser 90 95
       Leu Gly Ser Leu Val Pro Tyr Asp Phe Asp Val Ser His Cys Asp Phe 100 105 110
```

Ile Val Thr Gly Lys Asp Glu Asp Asp Ile Asn Val Trp Thr Val Met

75

| | Thr | Ala | 115 Phe | Glu | Tyr | Vāl | Gly | 120 Gly | Thr | Ile | Gly | Asn | 125 Gly | Val | Leu | Leu |
|-----|-------------|------------|------------|----------------|---------------|-------------|------------|-------------|------------|--------------|------------|------------|------------|------------|-----------------|------------|
| _ | | 130 His | | | | | 135 | | | | | 140 | | | | |
| 5 | 145 Lys | Агр | Phe | Pro | Asn | 150 Asn | Arg | Leu | Net | Glγ | 155 Val | Alā | Ile | Ala | Ser | 160 Asn |
| | Tyr | Arg | Ala | Pro | 165 Ser | Pro | Tyr | Gly | Leu | 170 Gly | Gly | Asp | Pro | Phe | 175 Ala | Leu |
| 10 | | ∨a1 | | 180 | | | | | 185 | _ | _ | | | 190 | | |
| | | Phe | 195 | | | | | 200 | | _ | | | 205 | | | _ |
| | | 210 Arg | | | | | 215 | | | | | 220 | | | | _ |
| 15 | 225 | Thr | | | | 230 | | | | - | 235 | | | | | 240 |
| | Vāl | Phe | Glu | Het | 245 Asn | Lys | Asn | Leu | Asp | 250 Gly | Phe | Āsp | Ile | Gl y | 255 Phe | Ile |
| 20 | Ser | Asn | Phe | 260 Val | Asp | Tyr | Агр | Pro | 265 Arg | Tyr | Ala | Trp | Ser | 270 Glu | Pro | Ile |
| | I]. e | Ile | 275 Glu | Glu | Asp | Сув | G1; | 280 Trp | Thr | Asp | Phe | Asn | 285 Pro | Leu | Gl _Y | Ala |
| | Leu | 290 Ser | Ile | Glu | Ile | Gln | 295 Het | Het | Leu | Asp | Asp | 300 Asn | Ser | Asp | Asn | Thr |
| 25 | 305 | Gly | | | | 310 | | | | | 315 | | | | | 320 |
| | | Val | | | 325 | | | | | 330 | | | | | 335 | |
| 30 | | Lys | | 340 | | | | | 345 | | | | | 350 | | |
| | Ala | Tyr | 355 Asp | Lys | Glu | Gly | Asp | 360 Arg | Туг | Leu | Thr | Thr | 365 Phe | Gln | Агр | His |
| 0.5 | Asn | 370 Leu | | | | | 375 | | | | | 380 | | | | |
| 35 | 385 T;r | Gly | Trp | Ser | | 390 Pro | туг | Val | Туг | | 395 Lys | Glu | Ala | Lys | Asp | 400 Lys |
| | Lγε | Arg | Arg | Arg | 405 Pro | Gln | Vāl | Ala | | 410 Asn | Pro | Thr | Asn | | 415 Lys | Ala |
| 40 | C7.2 | Trp | | 420 Trp | His | Thr | Arg | | 425 Ser | Pro | Tyr | Asp | | 430 Thr | Lys | Pro |
| | His | Pro 450 | 435 Thr | Pro | Val | Ile | | 440 Lys | His | Phe | Leu | | 445 Ser | Asp | Thr | Glu |
| 45 | Trp | Val | His | Ala | Leu | Asp 470 | 455 Val | Gly | Asp | Val | | 460 Gln | Lys | Glu | Gly | |
| 10 | | ras | Leu | Tyr | Pro 485 | - | Pro | Ala | Lys | | 475 T;r | Val | Leu | Il∈ | | 480 Leu |
| | Pro | Γε | Glu | Gly 500 | | His | Glu | Alā | Val 505 | 490 Val | T;r | Asp | Het | Gln 510 | 495 Gly | Arg |
| 50 | lle | Val | G1u 515 | | Val | Ser | Phe | Ser 520 | | Lys | Glu | Tjr | Lys 525 | | Asn | Val |
| | Gln | T;r 530 | | Ser | Lys | Gl 7. | Thr 535 | | Het | Leu | Lys | Val 540 | | Alā | Asp | Thr |
| 55 | Gl.u 545 | Tyr | Phe | Val | Glu | Lys 550 | | Ile | Vāl | Glu | | 340 | | | | |
| | | INFO | ORI4A'I | MOIT | FOR | | ID I | 10:39 | 0 | | | | | | | |
| | | | | QUENC | | | | | | | | | | | | |
| 60 | | | | 4) LE 3) Ti | | | | | acio | is | | | | | | |
| | | | (I |) TC | POLO | GY: | line | ear | | | | | | | | |
| 65 | | | | PECNI | | | _ | ein | | | | | | | | |
| | (| (iii) | НУІ | POTHE | TICA | AL: Y | ES | | | | | | | | | |
| 70 | | (vi) | | (GINA A) OF | | | | ohyro | mons | s gi | .ngiv | alis | • | | | |
| 70 | | (ix) | | TURE | | | | | | | | | | | | |
| | | | (<i>F</i> | 3) LC | MIE/P CATI | EY: ON 1 | misc | _fea 550 | ture | - | | | | | | |
| 75 | | (xi) | SEÇ | 2UENC | E DE | SCRI | FTIC | M: S | EQ I | ם ווכ |):390 | , | | | | |

| | Het 1 | Γ?.ε | Γεn | Γλε | Ile 5 | Alā | Leu | Arg | Leu | Leu 10 | Leu | Ala | Thr | Phe | Ala 15 | Ile |
|------------|----------|------|-----------|-----------|-------------------|-----|------|-----------|-----------|-----------|-----|-----|-----------|------------|-----------|-----|
| 5 | Val | Leu | Phe | Ser 20 | Pro | Leu | Ala | Lys | Ala 25 | Gln | Het | Asp | Ile | 30 GJ?. | Gly | Asp |
| | Asp | Val | Leu 35 | Ile | Glu | Thr | Met | Ser 40 | Thr | Leu | Ser | Glγ | Tyr 45 | Ser | Glu | Asp |
| | | 50 | | - | Met | | 55 | | | | | 60 | | | | |
| 10 | 65 | | | | Arg | 70 | _ | | | | 75 | | | | | 80 |
| | - | | | | Ala 85 | | | | _ | 90 | | | | | 95 | |
| 15 | | | | 100 | Phe | | | | 105 | | | | | 110 | | |
| | = | | 115 | | Asp | | | 120 | | | | | 125 | | | |
| 20 | _ | 130 | | | Thr | | 135 | | | | | 140 | | | | |
| 20 | 145 | | | | Asn | 150 | | | | | 155 | | | | | 160 |
| | | | | | Het 165 | | | | | 170 | | | - | | 175 | |
| 25 | | | | 180 | Leu | | | | 185 | | | | | 190 | | |
| | | | 195 | | Ser | | | 200 | | | | | 205 | | | |
| 20 | | 210 | | | His | | 215 | | | | | 220 | | | | |
| 30 | 225 | | | | Asn - | 230 | | _ | | | 235 | _ | | | | 240 |
| | | | _ | | Asn 245 | | = | | | 250 | _ | | | | 255 | |
| 3 5 | | | | 260 | Asp | | | | 265 | | | | | 270 | | |
| | | | 275 | | Arg | | | 280 | | | | | 285 | | | |
| 40 | | 290 | | | Thr | | 295 | | | | | 300 | | | | |
| 40 | 305 | | | | Leu | 310 | _ | | | _ | 315 | | | _ | _ | 320 |
| | | | | | 25 325 | | | | | 330 | | | | | 335 | |
| 4 5 | | | | 340 | Asn | | | | 345 | | | | | 350 | | |
| | | | 355 | _ | His | | | 360 | | | | | 365 | | | _ |
| 50 | | 370 | | | Tyr | | 375 | | | | | 380 | | | | |
| 50 | 385 | | | | με Γ.λ.ε | 390 | | | | | 395 | | | | | 400 |
| | | | | | Tyr 405 | | | | | 410 | | | | | 415 | |
| 55 | | | | 420 | | | | | 425 | _ | _ | | _ | 430 | | Pro |
| | | | 435 | | His | | | 440 | | | | | 445 | | | |
| 60 | | 450 | | | Asp | | 455 | | | | | 460 | | | | |
| 00 | 465 | | | _ | - | 470 | | | _ | | 475 | | | _ | | 480 |
| | | | | | Lys 485 Val | | | | | 490 | | | | | 495 | |
| 65 | | | | 500 | | | | | 505 | | _ | | | 510 | | _ |
| | | | 515 | | Het | | | 520 | - | | | | 525 | _ | | Ser |
| 70 | | 530 | | | Vāl | | 535 | vaı | * 9 1 | VIG | veh | 540 | OLU | 13.1 | 1112 | val |
| - = | 545 | -15 | | | | 550 | | | | | | | | | | |
| | 121 | THE | ORUAN | ייטדיי | FOR | SEO | י חד | 10.3 | 91 | | | | | | | |
| 75 | 1 - 1 | 2116 | - 14 1A | . 1011 | LOK | 220 | 101 | .0.3 | - 1 | | | | | | | |

75 (2) THEORYALTON FOR SEQ 10 NO. 391

```
(i) SEQUENCE CHARACTERISTICS:
                      (A) LENGTH: 390 amino acids
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
  5
             (ii) HOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
10
             (vi) ORIGINAL SOURCE:
                     (A) ORGANISH: Porphyromonas gingivalis
             (ix) FEATURE:
                     (A) NAME/KEY: misc_feature
15
                     (B) LOCATION 1...390
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:391
        Het Lys Arg Leu Leu Pro Phe Leu Leu Leu Ala Gly Leu Val Ala Val 1 5 10 15
20
        Gly Asn Val Ser Ala Gln Ser Pro Arg The Pro Gln Val Asp Val His 20 25 30
        Thr Arg Ile Ala Arg Asn Ala Arg Tyr Arg Leu Asp Lys Ile Ser Val 35 40 45
25
        Pro Asp Ser Arg Gln Ile Phe Asp Tyr Fhe Tyr Lys Glu Glu Thr Ile 50 \hspace{1cm} 60
        Pro Thr Lys Ile Gln Thr Thr Gly Gly Ala Ile Thr Ser Ile Asp 65 70 75 80
        Ser Leu Phe Tyr Glu Asp Asp Arg Leu Val Gln Val Arg Tyr Phe Asp
85 90 95
30
        Asn Asn Leu Glu Leu Lys Gln Ala Glu Lys Tyr Val Tyr Asp Gly Ser
100 105 110
        Lys Leu Val Leu Arg Glu Ile Arg Lys Ser Pro Thr Asp Glu Thr Pro
115 120 125

      Ile Lys Lys Val
      Ser Tyr His Tyr Leu Cys Gly Ser Asp Met Pro Phe 130
      135
      140

      Glu Ile Thr Thr Glu Met Ser Asp Gly Tyr Phe Glu Ser His Thr Leu 155
      160

35
        Asn Tyr Leu Asn Gly Lys Ile Ala Arg Ile Asp Ile Met Thr Gln Gln 165 170 175
40
        Asn Pro Ser Ala Glu Leu Ile Glu Thr Gly Arg Met Val Tyr Glu Phe
180 185 190
        Asp Ala Asn Asn Asp Ala Val Leu Leu Arg Asp Ser Val Phe Leu Pro 195 200 205
45
        Leu Gin Asn Lys Trp Val Glu Met Phe Thr His Arg Tyr Thr Tyr Asp 210 215 220
        Asn Lys His Asn Cys Ile Arg Trp Glu Gln Asp Glu Phe Gly Thr Leu
225 230 235 240
        Thr Leu Ala Asn Asn Phe Glu Tyr Asp Thr Thr Ile Pro Leu Ser Ser 245 250 255
50
        Val Leu Phe Pro Thr His Glu Glu Phe Phe Arg Pro Leu Leu Pro Asn
260 265 270
        Fhe Het Lys His Het Arg Thr Lys Gln Thr Tyr Phe Asn Asn Ser Gly 275 280 285
55
        Glu Gly Leu Ser Glu Val Cys Asp Tyr Asn Tyr Phe Tyr Thr Asp Met
290 295 300
       Gln Gly Asn Ala Leu Thr Asp Val Ala Val Asn Glu Ser Ile Lys Ile 305 310 315 320

Tyr Pro Arg Pro Ala Thr Asp Phe Leu Arg Ile Glu Gly Ser Gln Leu 325 335 335
60
       Leu Arg Leu Ser Leu Phe Asp Met Asn Gly Lys Leu Ile Arg Ala Thr 340 345 350

Glu Leu Thr Gly Asp Leu Ala Ile Ile Gly Val Ala Ser Leu Pro Arg 355 360 365
       Gly Thr Tyr Ile Ala Glu Ile Thr Ala Ala Asn Ser Lys Thr Ile Arg 370 375
        Ala Lys Val Ser Leu Arg
70
        (2) INFORMATION FOR SEQ ID NO: 392
               (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 411 amino acids
                     (B) TYPE: amino acid
75
                     (D) TOPOLOGY: linear
```

```
(ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
 5
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
10
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...411
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 392
15
       Net Arg Gln His Leu Ser Leu Phe Pro Phe Ile Leu Phe Leu Leu Leu
       Ala Phe Ser Tyr Val Gly Cys Arg Thr Val Arg Gln Thr Pro Lys Gln
20 25 30
       Ser Glu Arg Tyr Val Val Val Leu Ser Leu Asp Gly Phe Arg Pro Asp 35 40 45
20
       Tyr Thr Asp Arg Ala Arg Thr Pro Ala Leu Asp Arg Met Ala Gln Glu
50 55
       Gly Leu Ser Gly Ser Leu Gln Pro Cys Phe Pro Ser Leu Thr Phe Pro 65 \phantom{000}70\phantom{000}75\phantom{000} 80
25
       Asn His Tyr Ser Net Ala Thr Gly Leu Tyr Pro Asp His His Gly Ile 85 \\ 90 \\ 95
       Val Ala Ash Glu Phe Val Asp Ser Leu Leu Gly Ile Phe Arg Ile Ser
100 105 110
       Asp Arg Lys Ala Val Glu Thr Pro Gly Phe Trp Gly Gly Glu Pro Val
115 120 125
30
       Trp Asn Thr Ala Ala Arg Gln Gly Ile Arg Thr Gly Val Tyr Phe Trp 130 135 140
       Val Gly Ser Glu Thr Ala Val Asn Gly Asn Arg Pro Trp Arg Trp Lys
145 150 150 160
35
       Lys Phe Ser Ser Thr Val Pro Phe Arg Asp Arg Ala Asp Ser Val Ile
165 170 175
       Ala Trp Leu Gly Leu Pro Glu Lys Glu Arg Pro Arg Leu Leu Met Trp
180 185 190
       Tyr Ile Glu Glu Pro Asp Met Ile Gly His Ser Gln Thr Pro Glu Ser
195 200 205
40
       Pro Leu Thr Leu Ala Met Val Glu Arg Leu Asp Ser Val Val Gly Tyr 210 220
       Phe Arg Lys Arg Leu Asp Ser Leu Pro Ile Ala Ala Gln Thr Asp Phe 225 230 235 240
       Ile Ile Val Ser Asp His Gly Met Ala Thr T;r Glu Asn Glu Lys Cys 245 250 255
45
       Val Asn Leu Ser His Tyr Leu Pro Ala Asp Ser Phe Leu Tyr Met Ala
260 265 270
       Thr Gly Ala Phe Thr His Leu Tyr Pro Lys Pro Ser Tyr Thr Glu Arg
275 280 285
50
       Ala Tyr Glu Ile Leu Arg Ala Ile Pro His Ile Ser Val Tyr Arg Lys
, 290 295 300
       Gly Glu Val Pro Lys Arg Leu Arg Cys Gly Thr Asn Pro Arg Leu Gly 305 310 320
55
       Glu Leu Val Val Ile Pro Asp Ile Gly Ser Thr Val Phe Phe Ala Ile
325 330 335
       Asn Glu Asp Val Arg Pro Gly Ala Ala His Gly Tyr Asp Asn Gln Ala 340 345
       Pro Glu Het Arg Ala Leu Leu Arg Ala Val Gly Pro Asp Phe Arg Pro 355 360 365
60
       Gly Ser Arg Val Glu Asn Leu Pro Asn Ile Thr Ile Tyr Pro Leu Ile
370 375 380
       Cys Arg Leu Leu Gly Ile Glu Pro Ala Pro Asn Asp Ala Asp Glu Thr
385 390 395 400
65
       Leu Leu Asn Gl; Leu Ile Arg Asp Lys Arg Pro
       (2) INFORMATION FOR SEQ ID NO:393
70
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 246 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
75
            (ii) MOLECULE TYPE: protein
```

```
(iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
  5
                    (A) ORGANISH: Porphyromonas gingivalis
             (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...246
 10
             (mi) SEQUENCE DESCRIPTION: SEQ ID NO:393
       Met Lys Val Gly Leu Phe Ile Pro Cys Tyr Val Asn Ala Val Tyr Pro
                                                 10
 15
        Glu Val Gly Ile Ala Thr Tyr Lys Leu Leu Lys Ser Leu Asp Ile Asp 20 25 30
        Val Asp Tyr Pro Met Asp Gln Thr Cys Cys Gly Gln Pro Met Ala Asn
35 40
       Ala Gly Phe Glu Gln Lys Ala Gln Lys Leu Ala Leu Arg Phe Glu Glu 50 60
 20
       Leu Phe Glu Ser Tyr Asp Val Val Gly Pro Ser Ala Ser Cys Val
65 70 75 80
       Ala Phe Val Lys Glu Asn Tyr Asp His Ile Leu Arg Pro Thr Gly His
85 90 95
25
       Val Cys Lys Ser Ala Ala Lys Val Arg Asp Ile Cys Glu Phe Leu His 100 \hspace{1cm} 105 \hspace{1cm} 105 \hspace{1cm} 110 \hspace{1cm}
       Asp Asp Leu Lys Ile Thr Ser Leu Pro Ser Arg Phe Ala His Lys Val
       Ser Leu His Asn Ser Cys His Gly Val Arg Glu Leu His Leu Ser Thr
130 135 140
30
       Pro Ser Glu Val His Arg Pro Tyr His Asn Lys Val Arg Arg Leu Leu
145 150 155 160
       Glu Het Val Gln Gly Ile Glu Val Phe Glu Pro Lys Arg Ile Asp Glu
165 170 175
       Cys Cys Gly Phe Gly Gly Met Tyr Ser Val Glu Glu Pro Glu Val Ser
180 185 190
35
       180 185 190
Thr Cys Het Gly His Asp Lys Val Leu Asp His Ile Ser Thr Gly Ala
195 200 205
       Glu Tyr Ile Thr Gly Pro Asp Ser Ser Cys Leu Het His Net Gln Gly
210 215 220
Val Ile Asp Arg Glu Lys Leu Pro Ile Lys Thr Ile His Ala Val Glu
40
                            230
       Ile Leu Ala Ala Asn Leu
45
       (2) INFORMATION FOR SEQ ID NO: 394
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 246 amino acids
50
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
60
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...246
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:394
65
       Met Asp Ile Val Ser Met Ala Asp Lys Ala Leu Val Val Glu Met Arg
       Asp Val Thr Leu Cys Gln Glu Glu Asn Val Ile Phe Gln Asn Leu Asn
70
       Leu Thr Leu Ser Ala Gly Asp Phe Val Tyr Leu Ile Gly Ser Val Gly
                35
                                      40
       Ser Gly Lys Ser Thr Leu Leu Lys Ala Leu Tyr Ala Glu Val Pro Ile
50 60
       Ser Ala Gly Tyr Ala Arg Val Ile Asp Tyr Asp Leu Ala Lys Leu Lys
65 75 80
75
```



```
Arg Lys Gln Leu Pro Tyr Leu Arg Arg Asn Leu Gly Ile Val Phe Gln 85\, 90 \, 95
       Asp Phe Gln Leu Leu Asn Gly Arg Thr Val Ala Glu Asn Leu Asp Phe
100 105 110
       Val Leu Arg Ala Thr Asp Trp Lys Asn Arg Ala Asp Arg Glu Gln Arg
115 120 125
       Ile Glu Glu Val Leu Thr Arg Val Gly Net Ser Arg Lys Ala Tyr Lys
130 135 140
      Arg Pro His Glu Leu Ser Gly Gly Glu Gln Gln Arg Val Gly Ile Ala
145 150 150 160
10
      Arg Ala Leu Leu Ala Lys Pro Ala Leu Ile Leu Ala Asp Glu Pro Thr
165 170 175
      Gly Asn Leu Asp Ser Val Thr Gly Leu Gln Ile Ala Ser Leu Leu Tyr
180 190
      Glu Ile Ser Lys Gln Gly Thr Ala Val Leu Het Ser Thr His Asn Ser
195 200 205
15
               195
       Ser Leu Leu Ser His Leu Pro Ala Arg Thr Leu Ala Val Arg Lys Asn
210 215 220
       Gly Asp Ala Ser Ser Leu Val Glu Leu Ser Ala Asp Ala Val Ser Arg
20
       Lys Asn Thr Glu Ile Asp
       (2) INFORMATION FOR SEQ ID NO:395
25
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 241 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
30
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
35
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAMAE/KEY: misc_feature (B) LOCATION 1...241
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:395
      Met Ala Asp Lys Ala Leu Val Val Glu Met Arg Asp Val Thr Leu Cys
45
                                               1.0
       Gln Glu Glu Asn Val Ile Phe Gln Asn Leu Asn Leu Thr Leu Ser Ala
                   20
                                          25
       Gly Asp Phe Val Tyr Leu Ile Gly Ser Val Gly Ser Gly Lys Ser Thr 35 40
50
      Leu Leu Lys Ala Leu Tyr Ala Glu Val Pro Ile Ser Ala Gly Tyr Ala 50 60
      Arg Val Ile Asp Tyr Asp Leu Ala Lys Leu Lys Arg Lys Gln Leu Pro 65 70 80
      Tyr Leu Arg Arg Asn Leu Gly Ile Val Phe Gln Asp Phe Gln Leu Leu 85 90 95
55
      Asn Gly Arg Thr Val Ala Glu Asn Leu Asp Phe Val Leu Arg Ala Thr
100 105 110
      Asp Trp Lys Asn Arg Ala Asp Arg Glu Glu Glu Glu Val Leu 115 120 125
60
      Thr Arg Val Gly Het Ser Arg Lys Ala Tyr Lys Arg Pro His Glu Leu 130 140
      Ser Gly Gly Glu Gln Gln Arg Val Gly Ile Ala Arg Ala Leu Leu Ala
145 150 155 160
      Lys Pro Ala Leu Ile Leu Ala Asp Glu Pro Thr Gly Asn Leu Asp Ser
165 170
65
      Val Thr Gly Leu Gin Ile Ala Ser Leu Leu Tyr Glu Ile Ser Lys Gln
180 185 190
      Gly Thr Ala Val Leu Met Ser Thr His Asn Ser Ser Leu Leu Ser His
195 200 205
70
      Leu Pro Ala Arg Thr Leu Ala Val Arg Lys Asn Gly Asp Ala Ser Ser 210 215
       Leu Val Glu Leu Ser Ala Asp Ala Val Ser Arg Lys Asn Thr Glu Ile
                             230
       Asp
75
```

```
(2) INFORMATION FOR SEQ ID NO:396
            (i) SEQUENCE CHARACTERISTICS:
 5
                 (A) LENGTH: 232 amino acids
                  (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
10
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                 (A) ORGAHISH: Porphyromonas gingivalis
15
           (in) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...2\overline{3}2
20
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:396
      Het Arg Asp Val Thr Leu Cys Gln Glu Glu Asn Val Ile Phe Gln Asn
      Leu Asn Leu Thr Leu Ser Ala Gly Asp Phe Val Tyr Leu Ile Gly Ser
25 30
25
      Val Gly Ser Gly Lys Ser Thr Leu Leu Lys Ala Leu Tyr Ala Glu Val 45
       Pro Ile Ser Ala Gly Tyr Ala Arg Val Ile Asp Tyr Asp Leu Ala Lys
50 55
30
      Leu Lys Arg Lys Gln Leu Pro Tyr Leu Arg Arg Asn Leu Gly Ile Val
65 70 75 80
      Phe Gln Asp Phe Gln Leu Leu Asn Gly Arg Thr Val Ala Glu Asn Leu
85 90 95
      Asp Phe Val Leu Arg Ala Thr Asp Trp Lys Asn Arg Ala Asp Arg Glu
100 105 110
35
      Gln Arg Ile Glu Glu Val Leu Thr Arg Val Gly Het Ser Arg Lys Ala
115 120 125
       Tyr Lys Arg Pro His Glu Leu Ser Gly Gly Glu Gln Gln Arg Val Gly
130 135 140
40
      Ile Ala Arg Ala Leu Leu Ala Lys Pro Ala Leu Ile Leu Ala Asp Glu
145 150 155 160
      Pro Thr Gly Asn Leu Asp Ser Val Thr Gly Leu Gln Ile Ala Ser Leu
165 170 175
      Leu Tyr Glu Ile Ser Lys Gln Gly Thr Ala Val Leu Het Ser Thr His
180 185 190
45
                                                       190
       Asn Ser Ser Leu Leu Ser His Leu Pro Ala Arg Thr Leu Ala Val Arg
              195
                                  200
                                                        205
       Lys Asn Gly Asp Ala Ser Ser Leu Val Glu Leu Ser Ala Asp Ala Val
          210
                              215
50
       Ser Arg Lys Asn Thr Glu Ile Asp
      225
       (2) INFORMATION FOR SEQ ID NO:397
55
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 219 amino acids
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
60
          (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
65
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...219
70
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:397
      Het Ile Glu Ile Ser Asn Leu Thr Lys Val Phe Arg Thr Glu Glu Ile
                                             10
75
      Glu Thr Val Ala Leu Asp Gly Val Ser Leu Lys Val Asp Lys Gly Glu
```

| | Phe Ile | 35 | | | | | 40 | | | | | 45 | | | |
|--|---|---|---|--|---|--|---|--|---|---|--|--|---|---|---|
| 5 | Asn Ile 50 | Leu | Gly | Leu | Leu | Asp 55 | Asn | Pro | Thr | Ser | Gl; 60 | Ile | Tyr | Lys | Leu |
| | Asp Gly 65 | | | | 70 | | | | | 75 | | | | | 80 |
| 40 | Arg Lys | | | 85 | | | | | 90 | | | | | 95 | |
| 10 | Glu Het | | 100 | | | | | 195 | | | | | 110 | | _ |
| | Val Lys | Ala 115 | Ser | Glu | Arg | Γ ' .ε | Glu 120 | Arg | Val | Glu | Glu | Ala 125 | Leu | Arg | Lys |
| 15 | Net Ser 130 | | | | _ | 135 | - | | | | 140 | | | | - |
| | Gl; Gln 145 | | | | 150 | | | | | 155 | | | | | 160 |
| 20 | Γλε Γen | | | 165 | | | | | 170 | | | | | 175 | |
| 20 | Gly Ala | | 180 | | | | | 185 | | | | | 190 | | |
| | Thr Ile | 195 | | | | | 200 | | | - | Ala | Arg 205 | Ser | Ala | Gly |
| 25 | Arg Ile 210 | Il∈ | Asn | Leu | Phe | Asp 215 | Gly | Lys | Ile | Arg | | | | | |
| | (2) INFO | ORHAT | поп | FOR | SEQ | ID 1 | 10:39 | 98 | | | | | | | |
| 30 | (i) |) SEÇ | | | IARAG | | | | 10 | | | | | | |
| | | (E | 3) Ti | PE: | amir OGY: | no ac | cid | 401 | • | | | | | | |
| | (i i) | 1101 | | | | | | | | | | | | | |
| 35 | (iii) | НҮН | POT'HE | CTICA | AL: Y | ŒS | | | | | | | | | |
| | (vi) | ORI | GIN | AL SO | OURCE | :: | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| 40 | | (F | | | SM: | Porp | ohyro | mona | as gi | .ngi\ | alis | = | | | |
| 40 | | (<i>F</i> FE <i>F</i> (<i>F</i> | ATURE A) NA | E: MHE/F | ŒY: | misc | _fea | | | .ngi\ | /alis | Ē | | | |
| | (ix) | (<i>F</i> (<i>F</i> (<i>F</i> | ATURE A) 144 B) LO | E: MIE/I DCATI | ŒY: [ON] | misc | c_fea 595 | ature | 9 | | | Ē | | | |
| 40 45 | (ix) (xi) | () FEA () (E | ATURE A) NA B) LO QUENO | E: NIE/F DCATI CE DE | ŒY: [ON] [SCR] | miso | e_fea 595 ON: S | ature | E IIC |): 398 | 3 | | Gl v | V-1 | Tlo |
| | (ix) (xi) Het Lys | (A) FEA (A) (E) (B) SEQ | ATURE A) HA B) LO QUEHO Phe | E: NIE/F DCATI CE DE Phe 5 | (EY: ION 1 ESCRI Lys | miso [PTIO | c_fea 595 ON: S | EEQ : | E ID NO Ala 10 |):398 Ser | } Ile | Leu | _ | 15 | |
| | (ix) (xi) Het Lys 1 Thr Ala | (A) FEA (A) (B) Glu Gly | ATURE A) NA B) LO QUENC Phe Ile 20 | E: AME/P DCATI CE DE Phe 5 Ile | (EY: ION I ISCRI Lys Leu | miso L IPTIO Het Phe | c_fea 595 ON: S Phe Cys | EQ The | D HO Ala 10 Phe | D:398 Ser Leu |] Ile Phe | Leu Ile | Phe 30 | 15 Phe | Gly |
| 45 | (ix) (xi) Het Lys Thr Ala Ile Val | (A) FEA (A) (E) SEQ Glu Gly Ala 35 | ATURE A) HA B) LO QUEHO Phe Ile 20 Gly | E: AME/F DCATI CE DE Phe 5 Ile Ile | ŒY: [ON] ESCR] Lys Leu Ala | miso IPTIO Het Phe Ser | e_fea 595 ON: S Phe Cys Lys 40 | EEQ The Ile 25 Ala | Ala 10 Phe | D:398 Ser Leu Gl; | Ile Phe Gly | Leu Ile Thr | Phe 30 Ile | 15 Phe Pro | Gly Lys |
| 45 | (ix) (xi) Het Lys 1 Thr Ala Ile Val Ile Glu 50 | Glu Gly Ala 35 Ala | ATURE A) NA B) LO QUENO Phe Ile 20 Gly Asn | E: NHE/H DCATI CE DE Phe 5 Ile Ile Ser | (EY: ION 1 ESCRI Lys Leu Ala Ile | miso IPTIO Het Phe Ser Leu 55 | c_fea 595 ON: S Phe Cys Lys 40 His | EEQ The Ile 25 Ala | Ala 10 Phe Thr | Ser Leu Gl; | Ile Phe Gly Ser | Leu Ile Thr 45 Ser | Phe 30 Ile Phe | 15 Phe Pro | Gly Lys Glu |
| 4 5 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 | Glu Gly Ala 35 Ala Ser | ATURE A) HA B) LO DUEHO Phe Ile 20 Gly Asn | E: AME/F DCATI CE DE Phe 5 Ile Ile Ser Asn | CEY: ION 1 ESCRI Lys Leu Ala Ile Pro 70 | miso i IPTIO Het Phe Ser Leu 55 Trp | Phe Cys Lys 40 His | EQ The Ile 25 Ala Ile Met | Ala 10 Phe Thr Unk | Ser Leu Gly Asn Thr | Ile Phe Gly Ser 60 Gly | Leu Ile Thr 45 Ser Lys | Phe 30 Ile Phe Asp | 15 Phe Pro Pro Glu | Gly Lys Glu Ser 80 |
| 4 5 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser | Glu Gly Ala 35 Ala Ser Leu | ATURE A) HA B) LO QUEHO Phe Ile 20 Gly Asn Ala Ser | E: AME/FOCATION CE DE Phe 5 Ile Ile Ser Asn 61n 85 | ŒY: [ON] ESCRI Lys Leu Ala Ile Pro 70 Ala | miso is IPTIO Het Phe Ser Leu 55 Trp | e fea 595 Phe Cys Lys 40 His Ser | EEQ The Ile 25 Ala Ile Met Ala | Ala 10 Phe Thr Unk Leu 11e 90 | Ser Leu Gly Asn Thr 75 Gly | Ile Phe Gly Ser 60 Gly | Leu Ile Thr 45 Ser Lys Ala | Phe 30 Ile Phe Asp | 15 Phe Pro Pro Glu Agn 95 | Gly Lys Glu Ser 80 Asn |
| 4 5 | (ix) (xi) Het Lys 1 Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn | (A) FEAR (A) FEAR (A) SEQ Glu Gly Ala 35 Ala Ser Leu Ile | ATUREAN NA B) LO QUENO Phe Ile 20 Gly Asn Ala Ser Thr | E: CE DE Phe 5 Ile Ile Ser Asn Gln 85 Gly | CEY: ION 1 ESCRI Lys Leu Ala Ile Pro 70 Ala Ile | miso L IPTIO Het Phe Ser Leu 55 Trp Val | c_fea 595 Phe Cys Lys 40 His Ser Glu Leu | EQ The Ile 25 Ala Ile Met Ala Asp 105 | Ala 10 Phe Thr Unk Leu Ile 90 Leu | Ser Leu Gly Asn Thr 75 Gly Asp | Ile Phe Gly Ser 60 Gly Gln Asn | Leu Ile Thr 45 Ser Lys Ala | Phe 30 Ile Phe Asp Lys Ser 110 | 15 Phe Pro Pro Glu Agn 95 Val | Gly Lys Glu Ser 80 Asn |
| 45 50 55 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn Het Ala | (A) FEA (A) (B) (B) (B) (B) (B) (B) (B) (B) (B) (B | ATURE A) NA A) LO Phe Ile 20 Gly Asn Ala Ser Thr 100 Ala | E: NIE/F NIE/F Phe 5 Ile Ile Ser Asn Gln 85 Gly | CEY: CON 1 Lys Leu Ala Ile Pro 70 Ala Ile Glu | misc i | C_fea 95 Phe Cys 40 His Ser Glu Leu Arg 120 | Phe Ile 25 Ala Ile Met Ala Asp 105 Arg | Ala ala | Ser Leu Gly Asn Thr 75 Gly Asp Leu | Ile Phe Gly Ser 60 Gly Gln Asn | Leu Ile Thr 45 Ser Lys Ala Leu Asp | Phe 30 Ile Phe Asp Lys Ser 110 Phe | 15 Phe Pro Pro Glu Asn 95 Val | Gly Lys Glu Ser 80 Asn Gly |
| 45 50 55 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn Het Ala Ser Gly 130 | (A) FEA. (A) FEA. (A) (B) SEQ (Glu Gly A) Ala 35 Ala Ser Leu Ile Ser 115 Lys | ATURE ATURE AND MA Phe Ile 20 Gly Asn Ala Ser Thr 100 Ala | E: NAME/F OCATI Phe F Ile Ile Ser Asn Gln 85 Gly Glu Val | Lys Leu Ala Ile Pro 70 Ala Ile Glu Val | misc i IPTIC Het Phe Ser Leu 55 Trp Val Phe Leu Ser 135 | Phe Cys Lys 40 His Ser Glu Leu Arg 120 Tyr | Phe Ile 25 Ala Ile Met Ala Asp 105 Arg | Ala 10 IIC Phe Thr Unk Leu IIe 90 Leu Ala Asp | Ser Leu Gly Asn Thr 75 Gly Asp Leu | Ile Phe Gly Ser G0 Gly Gln Asn Gln Tyr | Leu Ile Thr 45 Ser Lys Ala Leu Asp 125 Thr | Phe 30 Ile Phe Asp Lys Ser 110 Phe Gln | 15 Phe Pro Glu Asn 95 Val Lys | Gly Lys Glu Ser 80 Asn Gly Met |
| 45 50 55 60 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn Ilet Ala Ser Gly 130 Tyr Tyr | (A) FEA (A) (B) (B) (C) (B) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C | ATURE ATURE B) LC Phe Ile 20 Gly Asn Ala Ser Thr 100 Ala Phe | E: NIE/F OCATI TE DE Phe 5 Ile Ile Ser Asn 6In 85 Gly Glu Val | KEY: ION 1 LSCSCRI Lys Leu Ala Ile Pro 70 Ala Ile Glu Val Ile 150 | misc i IPTIO Het Phe Ser Leu 55 Trp Val Phe Leu Ser 135 Ala | fes 595 Phe Cys 40 His Ser Glu Leu Arg 120 Tyr | Phe Ile 25 Ala Ile Met Ala Asp 105 Arg Ala Lys | Ala 10 Phe Thr Unk Leu Ala Asp Leu | Ser Leu Gly Asn Thr 75 Gly Asp Leu Arg Tyr 155 | Ile Phe Gly Ser 60 Gly Gln Asn Gln Tyr 140 Leu | Leu Ile Thr 45 Ser Lys Ala Leu Asp 125 Thr | Phe 30 Ile Phe Asp Lys Ser 110 Phe Gin Pro | 15 Phe Pro Glu Asn 95 Val Lys Lys | Gly Lys Glu Ser 80 Asn Gly Met Gly Gly 160 |
| 4550556065 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn Het Ala Ser Gly 130 Tyr Tyr | (A) FELA (A) (B) (B) (B) (B) (B) (B) (B) (B) (B) (B | ATURE ATURE AND LO QUENC Phe 11e 20 Gly Asn Ala Ser Thr 10c Ala Phe Ser Leu | E: NME/F OCATI TE DE Phe 5 Ile Ile Ser Asn 6In 85 Gly Glu Val Ser Ile | Lys Leu Ala Ile Pro Ala Ile Glu Val Ile 150 Gly | misc IPTIC Het Phe Ser Leu 55 Trp Val Phe Leu Ser 135 Ala | Cys Phe Cys 40 His Ser Glu Leu Arg 120 Tyr Asp | Phe Ile 25 Ala Ile Met Ala Asp 105 Arg Ala Lys | ID IIC Ala 10 Phe Thr Unk Leu Ala Asp Leu Gln 170 | Ser Leu Gly Asn Thr 75 Gly Asp Leu Arg Tyr 155 Thr | Ille Phe Gly Ser 60 Gly Gln Asn Gln Tyr 140 Leu | Leu Ile Thr 45 Ser Lys Ala Leu Asp 125 Thr Asn | Phe 30 Ile Phe Asp Lys Ser 110 Phe Gin Pro | 15 Phe Pro Pro Glu Asn 95 Val Lys Lys Lys | Gly Lys Glu Ser 80 Asn Gly Met Gly 160 Asp |
| 45 50 55 60 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn Het Ala Ser Gly 130 Tyr Tyr 145 Het Leu | (A) FEA (A) (B) (A) (B) (A) (B) (A) (B) (B) (B) (B) (B) (B) (B) (B) (B) (B | ATURE ATURE AND IMA Phe Ile 20 Gly Asn Ala Ser Thr 100 Ala Phe Leu Lye 180 | C: NIE/POCATI CE DE Phe 5 Ile Ile Ser Asn 6In 85 Gly Glu Val Ser Ile 165 Phe | KEY: ION 1 LSCRI Lys Leu Ala Ile Pro 70 Ala Ile Glu Val Ile 150 Gly | misc i iPTIO Het Phe Ser Leu 55 Trp Val Phe Leu Ser 135 Ala Ile | fes 595 Phe Cys 40 His Ser Glu Leu Arg 120 Tyr Asp | Phe Ile 25 Ala Ile Met Ala Asp 105 Arg Ala Lys Thr | Ala 10 Phe Thr Unk Leu Ala Asp Leu Gln 170 Glu | Ser Leu Gly Asn Thr 75 Gly Asp Leu Arg Tyr 155 Thr | Ile Phe Gly Ser 60 Gly Gln Asn Gln Tyr 140 Leu Het | Leu Ile Thr 45 Ser Lys Ala Leu Asp 125 Thr Asn Phe | Phe 30 Ille Phe Asp Lys Ser 110 Phe Gin Pro Tyr Val 190 | 15 Phe Pro Glu Asn 95 Val Lys Lys Lys Gly | Gly Lys Glu Ser 80 Asn Gly Met Gly 160 Asp |
| 4550556065 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn Het Ala Ser Gly 130 Tyr Tyr 145 Het Leu Ala Leu Tyr Lys Asn Arg | (A) FELA (A) (B) FELA (A) (B) SEC (B) Ala (B) | ATURE Phe Ile 20 Gly Asn Ala Ser Thr 100 Ala Phe Ser Leu Lys 180 Ala | E: NME/F OCATI TE DE Phe 5 Ile Ile Ser Asn 6In 85 Gly Val Ser Ile 165 Phe | KEY: ION 1 ESCRI Lys Leu Ala Ile Pro 70 Ala Ile Glu Val Ile 150 Gly Gly Glu | misc IPTIC Het Phe Ser Leu 55 Trp Val Phe Leu Ser 135 Ala Ile Val | Cys Phe Cys Lys 40 His Ser Glu Leu Arg 120 Tyr Asp Ala Lys | Phe Ile 25 Ala Ile Met Ala Asp 105 Arg Ala Lys Thr Met 185 Met | ID IIC Ala 10 Phe Thr Unk Leu Ala Asp Leu Gln 170 Glu Leu Leu | Ser Leu Gly Asn Thr 75 Gly Asp Leu Arg Tyr 155 Thr Ile | Ille Phe Gly Ser 60 Gly Gln Asn Gln Tyr 140 Leu Het | Leu Ile Thr 45 Ser Lys Ala Leu Asp 125 Thr Asn Phe Lys Met 205 | Phe 30 Ille Phe Asp Lys Ser 110 Phe Gin Pro Tyr Vall 190 Ser | 15 Phe Pro Glu Asn 95 Val Lys Lys Lys Gly Asp | Gly Lys Glu Ser 80 Asn Gly Het Gly Gl; Thr |
| 4550556065 | (ix) (xi) Het Lys Thr Ala Ile Val Ile Glu 50 Ile Val 65 Val Ser Pro Asn Het Ala Ser Gly 130 Tyr Tyr 145 Het Leu Ala Leu | (A) FEA. (A) (B) FEA. (A) (B) SEC (B) Ala 35 Ala Ser Leu Ile Ser 115 Lys Leu Gly Asp Ala 195 Glu | ATURE ATURE AND IMA Phe Ile 20 Gly Asn Ala Ser Thr 100 Ala Phe Ser Leu Lys 180 Ala Gln | C: NIE/POCATI CE DE Phe 5 Ile Ile Ser Asn Gin 85 Gly Glu Val Ser Ile 165 Phe Val | CEY: CON 1 CSCRI Lys Leu Ala Ile Pro 70 Ala Ile Glu Val Ile 150 Gly Gly Glu Thr | misc i | fes 595 Phe Cys 40 His Ser Glu Leu Arg 120 Tyr Asp Ala Lys Phe 200 Tyr | Phe Ile 25 Ala Ile Met Ala Asp Ala Lys Thr Met 185 Met Ile | Ala 10 Phe Thr Unk Leu Ala Asp Leu Ala Asp Leu Asn Asn | Ser Leu Gly Asn Thr 75 Gly Asp Leu Arg Tyr 155 Thr Ile Asn Gly | Ile Phe Gly Ser 60 Gly Gln Asn Gln Tyr 140 Leu Het Arg Leu 220 | Leu Ile Thr 45 Ser Lys Ala Leu Asp 125 Thr Asn Phe Lys Net 205 Trp | Phe 30 Ille Phe Asp Lys Ser 110 Phe Gln Pro Tyr Val 190 Ser Asp | 15 Phe Pro Glu Asn 95 Val Lys Lys Lys Lys Gly Asp | Gly Lys Glu Ser 80 Asn Gly Met Gly 160 Asp Thr Ala |

```
230
                                                     235
        Het Phe Ala Asp Lys Gly Glu Met Phe Gly Leu Ala Glu Lys Ala Val 245 255
        Glu Het Lys Leu Val Asp Glu Leu Ala Tyr Arg Thr Asp Val Glu Lys
260 265 270
  5
        Glu Leu Lys Lys Met Ser Gln Arg Gly Glu Lys Asp Glu Leu Arg Phe
275 280 285
        Val Ser Leu Ser Gln Val Leu Ala Asn Gly Pro Met Asn Lys Thr Lys
290 295 300
 10
        Gly Ser Arg Ile Ala Val Leu Phe Ala Glu Gly Glu Ile Thr Glu Glu 305 310 315
        Ile Ile Lys Lys Pro Phe Asp Thr Asp Gly Ser Ser Ile Thr Gln Glu
325 330 335
       Leu Ala Lys Glu Ile Lys Ala Ala Ala Asp Asp Asp Asp Ile Lys Ala 340 $345$
 15
        Val Val Leu Arg Val Asn Ser Pro Gly Gl; Ser Ala Phe Thr Ser Glu
355 360 365
       Gln Ile Trp Lys Gln Val Ala Asp Leu Lys Ala Lys Lys Pro Ile Val
370 375 380
20
       Val Ser Het Gly Asp Val Ala Ala Ser Gly Gly Tyr Tyr Ile Ala Cys
385 390 395 400
       Ala Ala Asn Ser Ile Val Ala Glu His Thr Thr Leu Thr Gly Ser Ile
405 415
       Gly Ile Phe Gly Met Phe Pro Asn Phe Ala Gly Val Ala Lys Lys Ile
420 430
25
       Gly Val Ash Net Asp Val Val Gln Thr Ser Lys Tyr Ala Asp Leu Gly
435 440
                                                           445
       Asn Thr Phe Ala Pro Met Thr Val Glu Asp Arg Ala Leu Ile Gln Arg
450 455 460
                                                       460
       Tyr Ile Glu Gln Gly Tyr Asp Leu Phe Leu Thr Arg Val Ser Glu Gly 470 475 480
30
       Arg Asn Arg Thr Lys Ala Gln Ile Asp Ser Ile Ala Gln Gly Arg Val
       Trp Leu Gly Asp Lys Ala Leu Ala Leu Gly Leu Val Asp Glu Leu Gly
500 510
35
       Gly Leu Asp Thr Ala Ile Lys Arg Ala Ala Lys Leu Ala Gln Leu Gly
515 520 525
       Gly Asn Tyr Ser Ile Glu Tyr Gly Lys Thr Lys Arg Asn Phe Phe Glu
530 535
40
       Glu Leu Leu Ser Ser Ser Ala Ala Asp Met Lys Ser Ala Ile Leu Ser 545 550 560
       Thr Ile Leu Ser Asp Pro Glu Ile Glu Val Leu Arg Glu Leu Arg Ser
565 570 575
       Het Pro Pro Arg Pro Ser Gly Ile Gln Ala Arg Leu Pro Tyr Tyr Phe
580 585 590
45
       Met Pro Tyr
       (2) INFORMATION FOR SEQ ID NO:399
50
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 589 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
55
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
60
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) HAME/KEY: misc_feature
65
                  (B) LOCATION 1...589
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:399
      Het Phe Phe Ala Ser Ile Leu Gly Val Ile Thr Ala Gly Ile Ile Leu
70
                                               10
       Phe Cys Ile Phe Leu Phe Ile Phe Phe Gly Ile Val Ala Gly Ile Ala
                    20
                                         25
       Ser Lys Ala Thr Gly Gly Thr Ile Pro Lys Ile Glu Ala Asn Ser Ile
               35
                                    40
75
       Leu His Ile Unk Asn Ser Ser Phe Pro Glu Ile Val Ser Ala Asn Pro
```

```
Trp Ser Het Leu Thr Gly Lys Asp Glu Ser Val Ser Leu Ser Gln Ala 65 70 75
        Val Glu Ala Ile Gly Gln Ala Lys Asn Asn Pro Asn Ile Thr Gly Ile
 5
                         85
                                                90
       Phe Leu Asp Leu Asp Asn Leu Ser Val Gly Het Ala Ser Ala Glu Glu
100 105 110
        Leu Arg Arg Ala Leu Gln Asp Phe Lys Met Ser Gly Lys Phe Val Val
115 120 125
10
       Ser T;r Ala Asp Arg Tyr Thr Gln L;s Gly Tyr Tyr Leu Ser Ser Ile 130 135 140
       Ala Asp Lys Leu Tyr Leu Asn Pro Lys Gly Het Leu Gly Leu Ile Gly 145 150 150 160
       Ile Ala Thr Gln Thr Met Phe Tyr Lys Asp Ala Leu Asp Lys Phe Gly
165 170 175
15
       Val Lys Met Glu Ile Phe Lys Val Gly Thr Tyr Lys Ala Ala Val Glu
180 185 190
        Pro Phe Het Leu Asn Arg Met Ser Asp Ala Asn Arg Glu Gln Ile Thr
195 200 205
20
       Thr Tyr Ile Asn Gly Leu Trp Asp Lys Ile Thr Ser Asp Ile Ala Glu
210 215 220
       Ser Arg Lys Thr Ala Net Asp Ser Val Lys Net Phe Ala Asp Lys Gly 225 230 235 240
       Glu Met Phe Gly Leu Ala Glu Lys Ala Val Glu Net Lys Leu Val Asp
245 250 255
25
       Glu Leu Ala Tyr Arg Thr Asp Val Glu Lys Glu Leu Lys Lys Met Ser
260 265 270
       Gln Arg Gly Glu Lys Asp Glu Leu Arg Phe Val Ser Leu Ser Gln Val
275 280 285
30
       Leu Ala Asn Gly Pro Het Asn Lys Thr Lys Gly Ser Arg Ile Ala Val
290 295 300
       Leu Phe Ala Glu Gly Glu Ile Thr Glu Glu Ile Ile Lys Lys Pro Phe 305 310 315
       Asp Thr Asp Gly Ser Ser Ile Thr Gln Glu Leu Ala Lys Glu Ile Lys
325 330
35
       Ala Ala Ala Asp Asp Asp Ile Lys Ala Val Val Leu Arg Val Asn 340 350
       Ser Pro Gly Gly Ser Ala Phe Thr Ser Glu Gln Ile Trp Lys Gln Val
355 360 365
40
       Ala Asp Leu Lys Ala Lys Lys Pro Ile Val Val Ser Met Gly Asp Val 370 380
       Ala Ala Ser Gly Gly Tyr Tyr Ile Ala Cys Ala Ala Asn Ser Ile Val
385 390 395 400
       Ala Glu His Thr Thr Leu Thr Gly Ser Ile Gly Ile Phe Gly Met Phe
405
410
415
45
       Pro Asn Phe Ala Gly Val Ala Lys Lys Ile Gly Val Asn Net Asp Val 420 425 430
       Val Gln Thr Ser Lys Tyr Ala Asp Leu Gly Asn Thr Phe Ala Pro Het
435 440 445
       Thr Val Glu Asp Arg Ala Leu Ile Gln Arg Tyr Ile Glu Gln Gly Tyr
450 455 460
50
       Asp Leu Phe Leu Thr Arg Val Ser Glu Gly Arg Asn Arg Thr Lys Ala 470 470 475
       Gln Ile Asp Ser Ile Ala Gln Gly Arg Val Trp Leu Gly Asp Lys Ala 485 490 425
55
       Leu Ala Leu Gly Leu Val Asp Glu Leu Gly Gly Leu Asp Thr Ala Ile
500 505 510
       Lys Arg Ala Ala Lys Leu Ala Gln Leu Gly Gly Asn Tyr Ser Ile Glu
515 520 525
60
       Tyr Gly Lys Thr Lys Arg Asn Phe Phe Glu Glu Leu Leu Ser Ser 530 535 540
       Ala Ala Asp Het Lys Ser Ala Ile Leu Ser Thr Ile Leu Ser Asp Pro 545 550 555 560
       Glu Ile Glu Val Leu Arg Glu Leu Arg Ser Net Pro Pro Arg Pro Ser 565 570 575
65
       Gly Ile Gln Ala Arg Leu Pro Tyr Tyr Phe Het Pro Tyr
580 585
       (2) INFORMATION FOR SEQ ID NO:400
70
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 313 amino acids
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
75
```

```
(ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
 5
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
10
                 (B) LOCATION 1...313
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 400
      Het Arg Ala Asn Ile Trp Gln Ile Leu Ser Val Ser Val Leu Phe
15
                                            10
      Phe Gly Thr Ala Ile Gly Gln Ala Gln Ser Arg Asn Arg Thr Tyr Glu
                                      25
                   20
      20
      Tyr Asn Ile Pro Ala Ser Ile Thr Ile Ala Gln Ala Leu Val Glu Thr
50 55 60
      Gly Ala Gly Ala Ser Thr Leu Ala Ser Val His Asn Asn His Phe Gly 65 70 75 80
      Ile Lys Cys His Lys Ser Trp Thr Gly Lys Arg Thr Tyr Arg Thr Asp 85 \\ 90 \\ 95
25
      Asp Ala Pro Ash Glu Cys Phe Arg Ser Tyr Ser Ala Ala Arg Glu Ser 100 105 110
      Tyr Glu Asp His Ser Arg Phe Leu Leu Gln Pro Arg Tyr Arg Pro Leu
115 120 125
30
      Fine Lys Leu Asp Arg Glu Asp Tyr Arg Gly Trp Ala Thr Gly Leu Gln 130 135 140
      Arg Cys Gly Tyr Ala Thr Asn Arg Gly Tyr Ala Asn Leu Leu Ile Lys
145 150 150 160
      Het Val Glu Leu Tyr Glu Leu Tyr Ala Leu Asp Arg Glu Lys Tyr Pro 165 170 175
35
      Ser Trp Phe His Lys Ser Tyr Pro Gly Ser Asn Lys Lys Ser His Gln 180 185 190
      Thr Thr Lys Gln Lys Gln Ser Gly Leu Lys His Glu Ala Tyr Phe Ser
195 200 205
      Tyr Gly Leu Leu Tyr Ile Ile Ala Lys Gln Gly Asp Thr Phe Asp Ser 210 215 220

Leu Ala Glu Glu Fhe Asp Met Arg Ala Ser Lys Leu Ala Lys Tyr Asn 225 230 230 235 240
40
      Asp Ala Pro Val Asp Phe Pro Ile Glu Lys Gly Asp Val Ile Tyr Leu
245 250 255
45
      Arg Val Gly Asp Ser Het His Ser Ile Ser Gln Arg Tyr Gly Ile Arg
275 280 285
50
      Met Lys Asn Leu Tyr Lys Leu Asn Asp Lys Asp Gly Glu Tyr Ile Pro
290 295 300
      Gln Glu Gly Asp Ile Leu Arg Leu Arg
55
       (2) INFORMATION FOR SEQ ID NO:401
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 523 amino acids
                 (B) TYPE: amino acid
60
                 (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
65
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
70
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...523
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:401
75
      Met Asp Gly Arg Arg Tyr Ser Asp Gly Leu His Gln Ala Ile Glu Ala
```

```
10
        Lys Glu His Val Lys Val Glu Ala Ala Thr Gln Thr Phe Ala Thr Ile
20 25 30
        Thr Leu Gln Asn Tyr Phe Arg Met Tyr His Lys Leu Ala Gly Met Thr 35 40
        Gly Thr Ala Glu Thr Glu Ala Gly Glu Leu Trp Asp Ile Tyr Lys Leu
50 60
        Asp Val Val Val Ile Pro Thr Asn Lys Pro Ile Ala Arg Lys Asp Het 65 70 75 80
       Asn Asp Arg Ile Tyr Lys Thr Ala Arg Glu Lys Tyr Ala Ala Val Ile
85 90
10
        Glu Glu Ile Val Arg Leu Val Glu Glu Gly Arg Pro Val Leu Val Gly
100 105 110
        Thr Thr Ser Val Glu Ile Ser Glu Leu Leu Ser Arg Met Leu Arg Leu 115 120 125
       Arg Gly Ile Gln His Asn Val Leu Asn Ala Lys Leu His Gln Lys Glu
130 135 140
        Ala Glu Ile Val Ala Gln Ala Gly Gln Lys Gly Thr Val Thr Ile Ala
145 150 150 160
20
       Thr Asn Met Ala Gly Arg Gly Thr Asp Ile Lys Leu Ser Ala Glu Val 165 170 175
        Lys Lys Ala Gly Gly Leu Ala Ile Ile Gly Thr Glu Arg His Glu Ser 180 185 190
       Arg Arg Val Asp Arg Gln Leu Arg Gly Arg Ser Gly Arg Gln Gly Asp 195 205

Pro Gly Ser Ser Ile Phe Tyr Val Ser Leu Glu Asp His Leu Het Arg 210 220
25
        Leu Phe Ala Thr Glu Lys Ile Ala Ser Leu Het Asp Arg Leu Gly Phe
225 230 235 240
30
        Lys Glu Gly Glu Val Leu Glu Asn Asn Het Leu Ser Lys Ser Val Glu
245 250 255
       Arg Ala Gln Lys Lys Val Glu Glu Asn Asn Phe Gly Ile Arg Lys His 260 265 270
        Leu Leu Glu Tyr Asp Asp Val Het Asn Ser Gln Arg Glu Val Ile Tyr
275 280 285
35
       Thr Arg Arg Arg His Ala Leu Net Gly Glu Arg Ile Gly Met Asp Val 290 295 300

Leu Asn Thr Ile Tyr Asp Val Cys Lys Ala Leu Ile Asp Asn Tyr Ala 305 310 315 320
40
        Glu Ala Asn Asp Phe Glu Gly Phe Lys Glu Asp Leu Met Arg Ala Leu
325 330 335
        Ala Ile Glu Ser Pro Ile Thr Gln Glu Ile Phe Arg Gly Lys Lys Ala 340 345 350
        Glu Glu Leu Thr Asp Met Leu Phe Asp Glu Ala Tyr Lys Ser Phe Gln
355 360 365
45
        Arg Lys Net Asp Leu Ile Ala Glu Val Ala His Pro Val Val His Gln
370 375 380
        Val Phe Glu Thr Gln Ala Ala Val Tyr Glu Arg Ile Leu Ile Pro Ile
385 390 395 400
50
        Thr Asp Gly Lys Arg Val Tyr Asn Ile Gly Cys Asn Leu Arg Glu Ala
405 416
        Asp Glu Thr Gln Gly Lys Ser Ile Ile Lys Glu Phe Glu Lys Ala Ile
420 425 430
        Val Leu His Thr Ile Asp Glu Ser Trp Lys Glu His Leu Arg Glu Met
435 440 445
55
        Asp Glu Leu Arg Asn Ser Val Gln Asn Ala Ser Tyr Glu Asn Lys Asp
450 455 460
        Pro Leu Leu Ile Tyr Lys Leu Glu Ser Tyr Glu Leu Phe Arg Lys Met
465 470 475 480
60
        Val Glu Ala Met Asn Arg Lys Thr Val Ala Ile Leu Met Arg Ala Arg
485 490 495
        Ile Pro Val Pro Glu Ala Pro Ser Gln Glu Glu Leu Glu His Arg Arg
500 505 510
        Gln Ile Glu Ile Arg His Ala Thr Gln Gln Arg
65
        (2) INFORMATION FOR SEQ ID NO:402
               (i) SEQUENCE CHARACTERISTICS:
70
                     (A) LENGTH: 375 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
```

(ii) MOLECULE TYPE: protein

75

```
(iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
  5
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...375
 10
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 402
        Het Asn Phe Leu Lys Lys Glu Pro Phe Lys Ile Phe Ser Met Ile Tyr 1 \phantom{\bigg|} 10 \phantom{\bigg|} 15
        Leu Leu Leu Asp Thr Ile Thr Asn Arg Ala Gly Thr Glu Arg Ala Val20 \\ 25 \\ 30
 15
        Ile Asn Leu Ala Asn Asn Leu His Ala Asn Gl; His Arg Val Ser Leu
35 40
        Val Ser Val Cys Thr Lys Glu Gly Glu Pro Ser Phe Gln Val Glu Lys
50 60
 20
        Gly Ile Glu Val His His Leu Gly Ile Arg Leu Tyr Gly Asn Ala Leu
65 70 80
       Ala Arg Lys Thr Val Tyr Phe Lys Ala Tyr Arg Arg Ile Lys Ala Leu
85 90 95
       Tyr Lys Lys Arg Glu Pro Val Leu Leu Ile Gly Thr Asn Ile Phe Ile 100 \ 105 \ 110
 25
       Asn Thr Ile Leu Ser Gln Ile Ser Asn Arg Gly Arg Ile Phe Thr Ile
115 120 125
       Gly Cys Glu His Ile Ser Tyr Asp Ile Ala Arg Pro Ile Thr Lys Arg
130 135 140
30
       Ile Arg Gly Phe Leu Tyr Ser Gly Leu Asp Ala Val Val Ala Leu Thr 145 150 155 160
       Lys Arg Asp Gln Gln Ser Phe Glu Ala Ile Leu Arg Gly Arg Ser Lys
165 170 175
       Ala Tyr Val Ile Pro Asn Gln Val Ser Phe Thr Thr Val Gln Arg Asp
180 185 190
35
       Ala Thr Thr His Lys Gln Het Leu Ala Ile Gly Arg Leu Thr Tyr Gln
195 200 205
       Lys Gly Phe Glu Phe Met Ile Glu Asp Ala Ser Arg Val Leu Arg Glu
210 215 220
40
       Arg Pro Asp Trp Lys Leu Ile Ile Val Gly Asp Gly Glu Asn Glu Ser 235 240
       Het Leu Arg Lys Glu Ile Ala Ser Arg Asn Met Glu Ser Gln Ile Glu
245 250 255
       Ile His Pro Ser Thr Pro Glu Ile Arg Lys Tyr Tyr Glu Ser Ser Ala
260 265 270
45
       Ile Tyr Leu Het Thr Ser Arg Phe Glu Gly Leu Pro Met Val Leu Leu 275 280 285
       Glu Ala Glu Ala Tyr Ala Leu Pro Ile Ile Ser Tyr Asp Cys Pro Thr
290 295 300
50
       Gly Pro Arg Glu Leu Ile Glu Asn Gly Arg Asn Gly Phe Leu Val Pro
305 310 315 320
       Het Glu Ala His Glu Asp Phe Ala Asp Lys Leu Arg Leu Leu Met Asp
325 330 335
       Asp Glu Thr Leu Arg Lys Lys Met Gly Gln Glu Ser Glu Leu Met Val
55
       Lys Ser Tyr Ser Pro Ala Asn Ile Tyr Glu Cys Trp Lys Leu Phe 355 360 365
       Val Glu Ile Gly Tyr Met Asn
60
       (2) INFORMATION FOR SEQ ID NO:403
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 362 amino acids
65
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
70
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Perphyromonas gingivalis
75
           (ix) FEATURE:
```

```
(B) LOCATION 1...362
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:403
 5
      Arg Ala Val Ile Asn Leu Ala Asn Asn Leu His Ala Asn Gly His Arg
20 25 30
      Val Ser Leu Val Ser Val Cys Thr Lys Glu Gl; Glu Pro Ser Phe Gln 35 40
10
      Val Glu Lys Gly Ile Glu Val His His Leu Gly Ile Arg Leu Tyr Gly
50 55 60
      Ash Ala Leu Ala Arg Lys Thr Val Tyr Phe Lys Ala Tyr Arg Arg Ile 65 70 75
15
      Lys Ala Leu Tyr Lys Lys Arg Glu Pro Val Leu Leu Ile Gly Thr Asn 85\, 90 95\,
      Ile Phe Ile Asn Thr Ile Leu Ser Gln Ile Ser Asn Arg Gly Arg Ile
100 105 110
20
      Phe Thr Ile Gly Cys Glu His Ile Ser Tyr Asp Ile Ala Arg Pro Ile 115 120 125
      Thr Lys Arg Ile Arg Gly Phe Leu Tyr Ser Gly Leu Asp Ala Val Val 130 135 140
      Ala Leu Thr Lys Arg Asp Gln Gln Ser Phe Glu Ala Ile Leu Arg Gly
145 150 155 160
Arg Ser Lys Ala Tyr Val Ile Pro Asn Gln Val Ser Phe Thr Thr Val
165 176 175
25
       Gln Arg Asp Ala Thr Thr His Lys Gln Net Leu Ala Ile Gly Arg Leu
180 185 190
      Thr Tyr Gln Lys Gly Phe Glu Phe Het Ile Glu Asp Ala Ser Arg Val
30
       Leu Arg Glu Arg Pro Asp Trp Lys Leu Ile Ile Val Gly Asp Gly Glu 210 215 220
       Ash Glu Ser Met Leu Arg Lys Glu Ile Ala Ser Arg Ash Met Glu Ser
225 230 235 240
35
      Gln lle Glu Ile His Pro Ser Thr Pro Glu Ile Arg Lys Tyr Tyr Glu
245 250 255
       Ser Ser Ala Ile Tyr Leu Met Thr Ser Arg Phe Glu Gly Leu Pro Met 260 270
40
       Val Leu Leu Glu Ala Glu Ala Tyr Ala Leu Pro Ile Ile Ser Tyr Asp
275 280 285
       Cys Pro Thr Gly Pro Arg Glu Leu Ile Glu Asn Gly Arg Asn Gly Phe
290 295 300
       Leu Val Pro Net Glu Ala His Glu Asp Phe Ala Asp Lys Leu Arg Leu
305 310 315 320
45
       Leu Het Asp Asp Glu Thr Leu Arg Lys Lys Het Gly Gln Glu Ser Glu 325 330 335
       Leu Met Val Lys Ser Tyr Ser Pro Ala Asn Ile Tyr Glu Cys Trp Lys 340 345 350
50
       Lys Leu Phe Val Glu Ile Gly Tyr Het Asn
       (2) INFORMATION FOR SEQ ID NO: 404
55
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 640 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
60
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
65
                 (A) ORGANISH: Perphyromenas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...640
70
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:404
       75
       Ser Val Leu Glu Gly Asn Glu Pro Ile Val Ile Thr Asn Ser Glu Gly
```

(A) NAME/KEY: misc feature

| | Lys | Arg | Thr 35 | 20 Thr | Pro | Ser | Val | Val 40 | 25 Ala | Phe | Val | Asp | Gly 45 | 30 Gly | Glu | Arg |
|------------|-----------|-----|------------|------------|-----------|------|------------|------------|------------|-----------|-----------|-----------|------------|------------|-----------|-----------|
| 5 | Γλε | Val | _ | Asp | Pro | Alā | Lys 55 | | Gln | Ala | Ile | Thr 60 | | Pro | Thr | Lys |
| Ü | Thr 65 | Ile | Tyr | Ser | Ile | L);s | Arg | Phe | llet | G13 | Glu 75 | | туг | Asp | Gln | Val 80 |
| | Ser | Arg | Glu | Val | Glu 85 | | Val | Pro | Phe | Lys 90 | | Val | Arg | Gly | Asp 95 | Asn |
| 10 | Asn | Thr | Pro | Arg 100 | | Asp | Ile | Asp | Gly 105 | Arg | Leu | Tyr | Thr | Pro 110 | Gln | Glu |
| | Ile | Ser | Ala 115 | llet | Ile | Leu | Gln | Lys 120 | Net | | Lys | Thr | Ala 125 | Glu | | T7.r |
| 15 | | 130 | | | | | Glu 135 | | | | | 140 | Pro | Ala | | |
| | 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| 00 | | | | | 165 | | Val | | | 170 | | | | | 175 | |
| 20 | | | | 180 | | | Asn | | 185 | | | | | 190 | | |
| | | | 195 | | | | Asp | 200 | | | | | 205 | _ | - | - |
| 25 | | 210 | | | | | Thr 215 | | | | | 220 | | | | |
| | 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| 20 | | | | | 245 | | Arg | | | 250 | | | | | 255 | |
| 30 | | | | 260 | | | Ala | | 265 | | | | | 270 | | |
| | | | 275 | | | | Tyr | 280 | | | | | 285 | | | |
| 35 | | 290 | | | | | Thr 295 | | | | | 300 | | | | |
| | 305 | | | | | 310 | Val | | | | 315 | | | | | 320 |
| 40 | | | | | 325 | | Asp | | | 330 | | | | | 335 | |
| 10 | | | | 340 | | | Ile Val | | 345 | | | | | 350 | | |
| | | | 355 | | | | Val | 360 | | | | | 365 | | | |
| 45 | | 370 | | | | | 375 Leu | | | | | 380 | | | | |
| | 385 | | | | | 390 | Asp | | | | 395 | | | | | 400 |
| 50 | | | | | 405 | | Ala | | | 410 | | | | | 415 | |
| | | | | 420 | | | Arg | | 425 | | | | | 430 | | |
| | | | 435 | | | | Gly | 440 | | | | | 445 | | | |
| 55 | | 450 | | | | | 455 Asp | | | | | 460 | | | | |
| | 465 | | | | | 470 | Thr | | | | 475 | | | | | 480 |
| 60 | | | | | 485 | | Asp | | | 490 | | | | | 495 | |
| | | | Alā | 500 | | | Λla | | 505 | | | | | 510 | | |
| C.F. | | lle | 515 | | | | Ser | 520 | | | | | 525 | | | |
| 65 | Lγε | 530 | | | | | 535 Phe | | | | | 540 | | | | |
| | 545 | | | | rλε | 550 | Lys | | | | 555 | | | | | 560 |
| 70 | | | | Thr | 565 | | Ala | | Leu | 570 | | | | Ser | 575 | |
| | | | Glu | 58 Q | | | Asn | Ala | 585 | | | | Gly | 590 | | |
| <i>7</i> 5 | | Gly | 595 | | | | Gly | 600 | | | | | 605 | | | |
| / 3 | | 610 | | | | | 615 | | | | | 620 | | | | |

PCT/AU98/01023

```
Pro Ser Asp Asp Lys Asn Val Thr Asp Val Asp Phe Glu Glu Val Lys
                              630
        (2) INFORMATION FOR SEQ ID HO: 405
  5
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 449 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 10
            (ii) MOLECULE TYPE: protein
           (iii) HTPOTHETICAL: YES
 15
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
 20
                   (B) LOCATION 1...449
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:405
       Het Arg Tyr Asp Leu Ala Ile Ile Gly Gly Gly Pro Ala Gly Tyr Thr
25
                                                10
       Ala Ala Glu Arg Ala Ala Lys Gly Gl; Leu Lys Thr Leu Leu Ile Glu 20 \\ 25 \\ 30
       Lys Asn Ala Leu Gly Gly Val Cys Leu Asn Glu Gly Cys Ile Pro Thr 35 40
30
       Lys Thr Leu Leu Tyr Ser Ala Lys Val Leu His Gln Ile Ala Thr Ala 50 $50$
       Ser Lys Tyr Ala Val Ser Gly Thr Ala Asp Gly Leu Asp Leu Gly Lys
65 70 75 80
       Val Ile Ala Arg Lys Gly Lys Ile Ile Arg Lys Leu Thr Ala Gly Ile
85 90 95
35
       Arg Ser Arg Leu Thr Glu Ala Gly Val Glu Het Val Thr Ala Glu Ala
       Thr Val Thr Gly Cys Asp Ala Asp Gly Ile Ile Gly Ile Thr Ala Gly
115 120 125
40
       Glu Ala Gln Tyr Lys Ala Ala Asn Leu Leu Cys Thr Gly Ser Glu
130 135 140
       Thr Phe Ile Pro Pro Ile Pro Gly Val Glu Gln Thr Glu Tyr Trp Thr
145 150 150
                                                  155
       Asn Arg Glu Ala Leu Gln Asn Lys Glu Ile Pro Thr Ser Leu Val Ile
165 170
45
                                             170
       Ile Gly Gl; Gly Val Ile Gly Met Glu Phe Ala Ser Phe Phe Asn Gly 180 185 190
       Ile Gly Thr Gln Val His Val Val Glu Het Leu Pro Glu Île Leu Asn
195 200 205
50
       Gly Ile Asp Pro Glu His Ala Ala Met Leu Arg Ala His Tyr Glu Lys
210 215 220
       Glu Gly Ile Lys Phe Tyr Leu Gly His Lys Val Thr Ser Val Arg Asn
225 230 235
       Gly Ala Val Thr Val Glu Tyr Glu Gly Glu Ser Lys Glu Ile Glu Gly 245 250 255
55
       Glu Arg Ile Leu Het Ser Val Gly Arg Arg Pro Val Leu Gln Gly Phe
260 265 270
       Glu Ser Leu Gly Leu Val Leu Ala Gly Lys Gly Val Lys Thr Asn Glu
275 280 285
60
       Arg Het Gln Thr Ser Leu Pro Asn Val Tyr Ala Ala Gly Asp Ile Thr
290 295 300
       Gly Phe Ser Leu Leu Ala His Thr Ala Val Arg Glu Ala Glu Val Ala
305 310 315 320
      Val Asp Gln Ile Leu Gly Lys Thr Asp Glu Thr Het Ser Tyr Arg Ala
325 330
65
       Val Pro Gl; Val Val Tyr Thr Asn Pro Glu Val Ala Gly Val Gly Glu
340 350
       Thr Glu Glu Ser Leu Arg Lys Ala Gly Arg Ala Tyr Thr Val Arg Arg
355 360 365
70
       Leu Pro Het Ala Phe Ser Gly Arg Phe Val Ala Glu Ash Glu Gln Gly
370 375 389
      Asn Gly Glu Cys Lys Leu Leu Leu Asp Glu Glu Asn Arg Leu Ile Gly 385 390 395
                                                 395
      Ala His Leu Ile Gly Asn Pro Ala Gly Glu Leu Ile Val Thr Ala Ala
75
                                               410
                                                                     415
```

```
Het Ala Ile Glu Thr Gly Met Thr Asp Arg Gln Ile Glu Arg Ile Ile
                   420
                                          425
        Phe Pro His Pro Thr Val Gly Glu Ile Leu Lys Glu Thr Leu Ala Gly
  5
       G1 y
        (2) INFORMATION FOR SEQ ID NO: 406
 10
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 941 amino acid\varepsilon
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
15
           (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
20
                  (A) ORGANISM: Porphyromenas gingivalis
            (ix) FEATURE:
                  (A) MAHE/KEY: misc_feature
                  (B) LOCATION 1...941
25
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:406
       Het Glu Leu Lys Arg Phe Leu Ser Leu Gly Leu Leu Val Gly Phe
                                              10
30
       Ile Pro Met Lys Leu Ser Ala Gln Gln Ala Gln Pro Leu Pro Thr Asp
20 25 30
       Pro Ala Val Arg Val Gly Lys Leu Asp Asn Gly Leu Thr Tyr Phe Ile
35 40
       Arg His Asn Glu Asn Pro Lys Asp Arg Ala Asp Phe Phe Ile Ala Gln 50 60
35
       Lys Val Gly Ser Ile Leu Glu Glu Asp Ser Gln Ser Gly Leu Ala His
65 70 75 80
       Phe Leu Glu His Het Ala Phe Asn Gly Thr Lys Asn Phe Pro Gly Lys
85 90 95
40
       Asn Leu Ile Asn Tyr Leu Glu Thr Ile Gly Val Arg Phe Gly Gln Asn
                   100
                                       105
       Leu Asn Ala Ser Thr Gly Phe Asp Lys Thr Glu Tyr Thr Ile Met Asp
115 120 125
       Val Pro Thr Thr Arg Gln Gly Ile Ile Asp Ser Cys Leu Leu Ile Leu
130 135 140
45
       His Asp Trp Ser Asn Asn Ile Thr Leu Asp Gly His Glu Ile Asp Glu
145 150 160
       Glu Arg Gly Val Ile Gln Glu Glu Trp Arg Ala Arg Arg Asp Ala Asn
165 170 175
50
       Leu Arg Het Phe Glu Ala Ile Leu Ala Lys Ala Met Pro Gly Asn Lys 180 185 190
       Tyr Ala Glu Arg Met Pro Ile Gly Leu Het Asp Val Val Leu Asn Phe
195 200 205
       Lys His Asp Glu Leu Arg Asn Tyr Tyr Lys Lys Trp Tyr Arg Pro Asp 210 215 220
55
       Leu Gln Gly Leu Val Ile Val Gly Asp Ile Asp Val Asp Tyr Val Glu
225 230 235 240
       Asn Lys Ile Lys Glu Leu Phe Lys Asp Val Pro Ala Pro Val Asn Pro
245 255
60
       Ala Glu Arg Ile Tyr Thr Pro Val Glu Asp Asn Asp Glu Pro Ile Val
260 265 270
       Ala Ile Ala Thr Asp Ala Glu Ala Thr Thr Thr Gln Leu Ser Ile Ser
275 280 285
       Phe Lys Ser Asp Pro Thr Pro Gln Glu Vai Arg Gly Ser Ile Phe Gly 290 295 300
65
       Leu Val Glu Asp Tyr Met Lys Gln Val Ile Thr Thr Ala Val Asn Glu
305 310 315
      Arg Leu Ser Glu Ile Thr His Lys Pro Asn Ala Pro Phe Leu Ser Ala
325 330 335
70
      Gl; Ala Phe Phe Ser Asn Phe Met Tyr Ile Thr Gln Thr Lys Asp Ala
340 345 350
       Phe Asn Phe Val Ala Thr Val Arg Glu Gly Glu Ala Glu Lys Ala Het 355 360 365
      Asn Ala Leu Val Ala Glu Ile Glu Ser Leu Arg Gln Phe Gly Ile Thr
370 375 380
75
                                 375
```

| | L√s 385 | Gly | Glu | Tyr | Asp | Arg 390 | Ala | Arg | Thr | Asn | Val 395 | Leu | Lys | Arg | Tyr | Glu 400 |
|------------|---------------------------|------------|------------|------------|------------|------------|------------|------------|--------------------------------|------------|------------|------------|------------|------------|------------|------------|
| | | Gln | Tyr | Asn | G1u 405 | | Asp | Lys | Arg | Lys 410 | | Asn | Ala | Tyr | Ala 415 | |
| 5 | Glu | Tyr | Ser | Thr 420 | | Phe | Thr | Asp | Gly 425 | | Tyr | Ile | Pro | Gly 430 | | Glu |
| | Val | Glu | T;r 435 | | Thr | Val | Asn | Ala 440 | | Ala | Pro | Gln | Val 445 | | Leu | Glu |
| 10 | Ala | Phe 450 | Asn | G1n | Ala | IJe | Ala 455 | | Met | 11e | Asp | Pro 460 | | Γλε | Asn | Ala |
| 10 | Val 465 | | Thr | Leu | Thr | Gly 470 | | Ser | $\mathrm{L}\lambda\varepsilon$ | Ala | Glu 475 | | Lys | Ile | Pro | Ser 480 |
| | | Ala | Asp | Phe | Leu 485 | | Ala | Phe | Lys | Ala 490 | | Arg | Gln | Gln | Lys 495 | |
| 15 | Glu | Ala | L/s | L;s 500 | | Glu | Val | Ser | Asp 505 | | ГЛг | Leu | Met | Glu 510 | | Ala |
| | Pro | Lys | Ala 515 | Gly | Lys | Ile | Val | Ser 520 | Glu | Lys | Lys | Asp | Gln 525 | Lys | Ph≘ | Gly |
| 20 | Thr | Thr 530 | Glu | Leu | Thr | Leu | Ser 535 | Asn | Glγ | Ile | Lys | Val 540 | Tyr | Leu | Lys | Lys |
| | 545 | | Phe | | | 550 | | | | | 555 | | | | | 560 |
| | Gly | Ile | Leu | Ser | Gly 565 | Γλε | His | Ala | Pro | Asn 570 | Gln | Ser | Val | Het | Asn 575 | Ser |
| 25 | Ph∈ | Het | Asn | Val 580 | G1 'y | Gl y | Leu | Gly | Asn 585 | Phe | Asp | Ala | Ile | Gln 590 | Leu | Ағр |
| | Γ ?. ε | Val | Leu 595 | Thr | Gl y | Arg | Ser | Ala 600 | Ser | Val | Ser | Pro | Ser 605 | Leu | Ser | Leu |
| 30 | | 610 | Glu | | | | 615 | _ | | | | 620 | - | | | |
| | 625 | | Gln | | | 630 | | | | | 635 | | | | | 640 |
| 25 | | | Phe | | 645 | | | | | 650 | | | | | 655 | |
| 35 | | | Ala | 660 | | | | | 665 | | = | | | 670 | | |
| | | _ | Gly 675 | | | | | 680 | _ | | | _ | 685 | | _ | |
| 4 0 | | 690 | Val | | _ | - | 695 | | | | | 700 | | | _ | |
| | 705 | _ | Ala | _ | _ | 710 | | | | | 715 | _ | | | - | 720 |
| 45 | | | llet | | 725 | | | | | 730 | | | | | 735 | |
| 10 | | | Arg Lys | 740 | | | | | 745 | | | | | 750 | | |
| | | | 755 Ile | | | | | 760 | | | | | 765 | | | |
| 50 | | 770 | Leu | | | | 775 | | | | | 780 | | | | |
| | 785 | | Thr | | | 790 | | | | | 795 | | | | | 800 |
| 55 | | | Gly | | 805 | | | | - | 810 | | - | | | 815 | |
| | | | Pro | 820 | | | | | 825 | | | | | 830 | | |
| | | | 835 Glu | | - | | | 840 | | | | | 845 | | | |
| 60 | Phe | 850 Lys | Lys | Thr | Ile | Glu | 855 Asn | Leu | Asn | Lys | Gln | 860 His | Lys | Glu | Ser | Leu |
| | 865 Arg | Glu | Asn | Arg | Phe | 870 Trp | Leu | Glu | Ala | Net | 875 Lys | Ala | Ser | Phe | Phe | 880 Glu |
| 65 | Glγ | Asn | Asp | | 885 Ile | Thr | Asp | Туr | | 890 Ser | Val | Leu | Asn | | 895 Leu | Thr |
| | Pro | Ala | Glu | 900 200 | Gln | Lys | Phe | | 905 Ala | Asp | Leu | Leu | | 910 Gln | Gln | Asn |
| 70 | Arā | Val 930 | 915 Val | Val | l4et | Met | Ala 935 | 920 Pro | Val | Al a | Lys | Ala 940 | 925 Gln | | | |
| | (2) | INFO | ORI-IA'I | NOI | FOR | SEQ | ID 1 | 10:40 | 7 | | | | | | | |
| | | | | | | | | | | | | | | | | |

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 684 amino acids

```
(B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
 5
           (iii) HTPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISH: Porphyromonas gingivalis
10
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...684
15
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:407
       Het Ser Lys Lys Gly Thr Ile Gly Val Thr Ser Asp Asn Ile Phe Pro
       Val Ile Lys Lys Phe Leu Tyr Ser Asp His Glu Ile Phe Leu Arg Glu 20 \phantom{-}25\phantom{+}30\phantom{+}
20
       Ile Val Ser Asn Ala Val Asp Ala Thr Gln Lys Leu Lys Thr Leu Thr 35 40 45
       Ser Val \widehat{\text{Gl}}\gamma Glu Phe Lys Gly Glu Thr Gly Asp Leu Arg Val Thr Val 50 55
       Ser Val Asp Glu Val Ala Arg Thr Ile Thr Val Ser Asp Arg Gl; Val
65 70 80
25
       Gly Met Thr Glu Glu Glu Val Glu Lys Tyr Ile Asn Gln Ile Ala Phe
85 90 95
       Ser Ser Ala Glu Glu Phe Leu Glu Lys Tyr Lys Asp Asp Lys Ala Ala
100 105 110
30
       Ile Ile Gly His Phe Gly Leu Gly Phe Tyr Ser Ala Phe Met Val Ser
115 120 125
       Glu Arg Val Asp Val Ile Thr Arg Ser Phe Arg Glu Asp Ala Thr Ala
130 135 140
35
       Val Lys Trp Ser Cys Asp Gly Ser Pro Glu Tyr Thr Leu Glu Pro Ala
145 150 155 160
       Asp Lys Ala Asp Arg Gly Thr Asp Ile Val Het His Ile Asp Glu Glu
165 170 175
       Asn Ser Glu Phe Leu Lys Lys Glu Lys Ile Glu Gly Leu Leu Gly Lys
180 185 190
40
       Tyr Cys Lys Phe Leu Thr Val Pro Ile Ile Phe Gly Lys Lys Gln Glu
195 200 205
       Trp Lys Asp Gly Lys Net Gln Asp Thr Asp Glu Asp Asn Gln Ile Asn 210 220
       Asp Thr His Pro Ala Trp Thr Lys Lys Pro Ala Asp Leu Lys Asp Glu
225 230 235 240
45
       Asp Tyr Lys Glu Phe Tyr Arg Ser Leu Tyr Pro Het Ser Glu Glu Pro
245 250 255
       Leu Phe Trp Ile His Leu Asn Val Asp Tyr Pro Phe Asn Leu Thr Gly 260 265 270
50
       Ile Leu Tyr Phe Pro Lys Ile Lys Asn Asn Leu Asp Leu Gln Arg Asn 275 280 285
       Lys Ile Gln Leu Tyr Cys Asn Gln Val Tyr Val Thr Asp Glu Val Gln 290 295 300

Gl; Ile Val Pro Asp Phe Leu Thr Leu Leu His Gl; Val Ile Asp Ser 305 310 315 320
55
       Pro Asp Ile Pro Leu Asn Val Ser Arg Ser Tyr Leu Gln Ser Asp Ala
325 330 335
       Asn Val Lys Lys Ile Ser Ser His Ile Thr Lys Lys Val Ala Asp Arg 340 345 350
60
       Leu Glu Glu Ile Phe Lys Asn Asp Arg Pro Thr Phe Glu Glu Lys Trp
355 360 365
       Asp Ser Leu Lys Leu Phe Val Glu Tyr Gly Het Leu Thr Asp Glu Lys 370 375 380
65
       Phe Tyr Glu Arg Ala Ala Lys Phe Phe Leu Phe Thr Asp Met Asp Gly 385 390 395 400
       His Lys Tyr Thr Fhe Asp Glu Tyr Arg Thr Leu Val Glu Gly Val Gln 405 415
       Thr Asp Lys Asp Gly Gln Val Val Tyr Leu Tyr Ala Thr Asp Lys His
420 425 430
70
       Gly Gln Tyr Ser His Val Lys Arg Ala Ser Asp Lys Gly Tyr Ser Val
435 440 445
       Het Leu Leu Asp Gly Gln Leu Asp Pro His Ile Val Ser Leu Leu Glu
450 455 460
75
       Gln Lys Leu Glu Lys Thr His Phe Val Arg Val Asp Ser Asp Thr Ile
```

```
465
                             470
       Asn Asn Leu Ile Arg Lys Glu Giu Arg Ala Glu Val Lys Leu Ser Asp
485 490 495
       Thr Glu Arg Ala Thr Leu Val Lys Leu Phe Glu Ala Arg Leu Pro Arg
 5
                   500
                                        505
                                                               510
       Asp Glu Lys Lys His Phe Asn Val Ala Phe Glu Ser Leu Gly Ala Glu
515 520 525
       Gly Glu Ala Ile Leu Ile Thr Gln Ala Glu Phe Het Arg Arg Met Arg 530 $535 $540
       Asp Net Ala Gln Leu Gln Pro Gly Net Ser Phe Tyr Gly Glu Leu Pro 545 550 560
10
       Asp Ser Tyr Asn Leu Val Leu Asn Thr Asp His Pro Leu Ile Asp Arg
565 570 575
       Val Leu Ser Gly Glu Lys Glu Ser Val Glu Pro Ser Leu Thr Glu Leu
580 585
15
       Arg Ala Lys Ile Ala Glu Leu Lys Ala Glu Glu Ala Lys Leu Leu Asp
595 600 605
       Glu Glu Lys Gly Lys Lys Pro Glu Glu Ile Pro Val Ala Thr Lys Glu 610 620
20
       Ala Lys Glu Asn Asn Ala Val Glu Gln Ala Lys Thr Glu Gly Ser Ile
625 630 635 640
       Asn Asp Gln Leu Thr Lys Tyr Ala Gln Asp Asn Glu Leu Ile Gly Gln
645
                                          650
       Leu Ile Asp Leu Ala Leu Leu Gly Ser Gly Leu Leu Thr Gly Glu Ala
660 665 670
25
       Leu Ala Glu Phe Ile Arg Arg Ser Gln Arg Leu Leu
       (2) INFORMATION FOR SEQ ID NO:408
30
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 464 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
35
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
40
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
45
                  (E) LOCATION 1...464
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:408
       Het Glu Lys Leu Ile Asp Ile Leu Val Val Asp Asp Asp Val Ala Val
50
       Cys Ala Ala Leu Arg Leu Val Leu Lys Arg Ala Gl; Tyr Asn Pro Val
20 25 30
       Ile Ala Asn Ser Pro Asp Glu Ala Leu Ser Ile Met Arg Asn Pro Asp
                                    40
55
       Gly Gly Cys Lys Pro Ala Val Ile Leu Met Asp Met Asn Phe Ser Leu
50 60
       Ser Thr Ser Gly Arg Glu Gly Leu Glu Leu Glu Lys Met Gln Ile
65 70 75 80
       Phe Thr Ser Cys Pro Val Ile Leu Het Thr Ala Trp Ala Ser Ile Pro
85 90
60
       Leu Ala Val Glu Gly Het Arg Leu Gly Ala Phe Asp Phe Ile Gly Lys
100 105 110
       Pro Trp Asp Asp Asp Arg Leu Leu Arg Thr Ile Asp Thr Ala Leu His
115 120 125
       Leu Ala Ala Pro Ser Ala Val Ala Asn Pro Ser Glu Gln Ser Asp Arg
130 135
65
       Asp Thr Ala Arg Gln Pro Lys Ala Thr Val Gln Glu Asn Asp Pro Cys
145 150 155 160
      Ala His Ile Ile Gly Arg Ser Asp Ala Ile Cys Lys Ile Lys Glu Arg
165 170 175
70
       Ile Arg Arg Ile Ala Pro Thr His Ala Ser Val Leu Ile Thr Gly Glu
                   180
                                      185
       Ser Gly Thr Gly Lys Glu Leu Ile Ala Glu Ala Leu His Arg Gly Ser
                                    200
                                                         205
75
       Lys Arg Ala Ser Ala Pro Phe Val Lys Val Asn Leu Gly Gly Ile Pro
```

```
210
                                    215
        Glu Ser Leu Phe Glu Ser Glu Leu Phe Gly His Lys Lys Gly Ala Phe 225 230 235 240
        Thr Asn Ala Phe Ser Asp Arg Lys Gly Arg Phe Glu Leu Ala Asp Gly 245 255
  5
        Gly Thr Ile Phe Leu Asp Glu Ile Gly Glu Leu Pro Val Gly Asn Gln
260 265 270
        Val Lys Leu Leu Arg Val Leu Gln Glu Gln Thr Phe Glu Pro Leu Gly 275 280 285
 10
        Glu Ser Val Ser His Arg Val Asp Ile Arg Val Val Ser Ala Thr Asn
290 295 300
        Ala Ser Leu Glu Arg Met Val Ala Glu Gly Arg Phe Arg Glu Asp Leu 305 310 315 320
        Tyr Tyr Arg Ile Asn Leu Ile His Leu His Leu Pro Pro Leu Arg Glu
325 330 335
 15
        Arg Gln Glu Asp Ile Gln Leu Leu Val Glu Ala Phe Ser Glu Ala Phe
340 345 350
        Ala Gln Ser Asn Gly Leu Pro His Ala Val Trp Ser Ala Glu Ala Het
355 360 365
 20
        Arg Arg Ile Cys Ala Met Pro Leu Pro Gly Asn Val Arg Glu Leu Lys
370 375 380
        Asn Val Val Glu Arg Thr Leu Leu Leu Ser Gly Ser Arg Glu Ile Ser
385 390 395 400
        Ala Arg Asp Val Ala Asp Phe Gly Ser Gln Val Thr Ala Ala Asp His
405 410 415
 25
        Ser Asp Glu Arg Ala Leu Thr Asp Met Glu Glu Ala Ala Ile Arg Glu 420 425 430
        Thr Leu Thr Lys Tyr Asn Gly Asn Val Ser Arg Ala Ala Arg Ala Leu 435 440
 30
        Gly Leu Ser Arg Ala Ala Leu Tyr Arg Arg Met Glu Lys Tyr Gly Leu
        (2) INFORMATION FOR SEQ ID NO:409
35
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 250 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
40
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) HAME/KEY: misc feature
                   (B) LOCATION 1...250
50
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 409
       Het Leu Lys Ile Lys Asn Leu His Ala Thr Val Gln Gly Lys Glu Ile
                                               10
55
       Leu Lys Gly Ile Asn Leu Glu Ile Asn Ala Gly Glu Ile His Ala Ile 20 25 30
       Het Gly Pro Asn Gly Ser Gly Lys Ser Thr Leu Ser Ser Val Leu Val
35 40 45
       Gly His Pro Ser Phe Glu Val Thr Glu Gly Glu Val Thr Phe Asn Gly
50 55
60
       Ile Asp Leu Leu Glu Leu Glu Pro Glu Glu Arg Ala His Leu Gly Leu
65 70 75 80
       Phe Leu Ser Phe Gln Tyr Pro Val Glu Ile Pro Gly Val Ser Het Val
       Asn Phe Het Arg Ala Ala Val Asn Glu His Arg Lys Ala He Gly Ala 100 105 110

Glu Pro Val Ser Ala Ser Asp Phe Leu Lys Het Het Arg Glu Lys Arg 115 120 125
65
      Ala Ile Val Glu Leu Asp Asn Lys Leu Ala Ser Arg Ser Val Asn Glu
130 135 140
70
      Gly Phe Ser Gly Gly Glu Lys Lys Arg Asn Glu Ile Phe Gln Het Ala
145 150 155 160
      Met Leu Glu Pro Lys Leu Ala Ile Leu Asp Glu Thr Asp Ser Gly Leu
165 170 175
      Asp Ile Asp Ala Leu Arg Ile Val Ala Gly Gly Val Asn Arg Leu Arg
```

```
180
                                              185

      Ser Pro Glu Asn Ala Ala Ile Val Ile Thr His Tyr Gln Arg Leu Leu

      195
      200
      205

      Glu Tyr Ile Lys Pro Asp Phe Val His Val Leu Tyr Lys Gly Arg Ile
      215
      220

       Val Lys Ser Gly Gly Ala Glu Leu Ala Leu Thr Leu Glu Glu Lys Gly
                             230
       Tyr Asp Trp Ile Lys Glu Glu Ile Gly Glu
10
        (2) INFORMATION FOR SEQ ID NO:410
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 461 amino acids
15
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
20
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                    (A) HAHE/KEY: misc_feature
                    (B) LOCATION 1...461
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:410
30
       Tyr Ala Lys Trp Gln Gly Asn Cys Asn Ala Cys Gly Glu Trp Asn Cys
20 25 30
35
       Ile Asp Glu Glu Lys Val Pro Ala Pro Ala Ser Gly Lys His Ala Ala 35 \phantom{-}40
       Lys Ser Phe Met Pro Arg Glu Gln Asp Asn Arg Pro Arg Leu Leu Gln 50 60
       Asp Val Glu Ser Gly Asp Glu Glu Arg Ile Arg Leu Gly Asp Glu Glu 65 70 75 80
40
       Phe Asp Arg Val Leu Gly Gly Gly Ile Val Lys Gly Ala Phe Val Leu 85 90 95
       Leu Gly Gly Glu Pro Gly Ile Gly Lys Ser Thr Leu Ile Leu Gln Thr
100 105 110
45
       Val Leu Arg Leu Pro Gln Leu Arg Thr Leu Tyr Val Ser Gly Glu Glu 115 120 125
       Ser Ala Arg Gln Leu Lys Met Arg Ala Glu Arg Leu Gly Gln Ala Met
130 135 140
       Asn Gly Cys Tyr Val Tyr Cys Glu Thr Asn 11e Glu Arg 11e Leu Ser
145 150 155 160
50
       Arg Ala Glu Glu Leu Thr Pro Asp Leu Leu Val Ile Asp Ser Ile Gln 165 170 175
       Thr Val Tyr Thr Glu Glu Met Glu Ser Ser Ala Gly Ser Val Gly Gln
180 185 190
55
       Ile Arg Glu Cys Ala Ala Leu Leu Leu Lys Tyr Cys Lys Thr Thr Gly 195 200 205
       Ile Pro Val Ile Val Ile Gly His Ile Thr Lys Glu Gly Ser Ile Ala
210 215 220
       Gly Pro Lys Val Leu Glu His Ile Val Asp Thr Val Leu Leu Phe Asp
225 230 235 240
60
       Gly Asp Lys His His Leu Tyr Arg Ile Leu Arg Gly Gln Lys Asn Arg 245 250 255

Tyr Gly Ser Thr Ser Glu Leu Gly Ile Tyr Glu Met Arg Gln Asp Gly 260 265 270
65
       Leu Arg Gly Val Glu Asn Pro Ser Glu His Leu Ile Thr Arg Asn Arg
275 280 285
       Glu Asp Leu Ser Gly Ile Ala Ile Ala Val Ala Met Glu Gly Ile Arg 290 295 300
                                                          300
       Pro Ile Leu Ile Glu Ala Gln Ala Leu Val Ser Ser Ala Ile Tyr Ala
305 310 315 320
70
       Asn Pro Gin Arg Ser Ala Thr Gly Phe Asp Ile Arg Arg Het Asn Het
325 330 335
       Leu Leu Ala Val Leu Glu Lys Arg Ala Gly Phe Lys Leu Ile Gln Lys 340 345 350
75
       Asp Val Phe Leu Asn Ile Ala Gly Gl; Ile L;s Ile Ala Asp Pro Ala
```

```
360
        Thr Asp Leu Ala Val Ile Ser Ala Val Leu Ala Ser Ser Leu Asp Ile
370 375 380
        Val Ile Pro Pro Ala Val Cys Met Thr Gly Glu Val Gly Leu Ser Gly 385 390 395 400
 5
        Glu Ile Arg Pro Val Ser Arg Ile Glu Gln Arg Ile Thr Glu Ala Arg
405 410 415
        Arg Ile Gly Phe Lys Glu Ile Leu Val Pro Ala Asp Asn Phe Arg Gln 420 430
10
        Glu Asp Ala Gly Arg Phe Gly Ile Arg Leu Val Pro Val Arg Lys Val
435 440 445
                                                            445
        Glu Glu Ala Phe Arg His Leu Phe Ser Lys Gly Arg Glu
450 455
15
        (2) INFORMATION FOR SEQ ID NO:411
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 271 amino acids
                    (B) TYPE: amino acid
20
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
25
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
30
                   (A) NAME/KET: misc_feature
                   (B) LOCATION 1...271
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:411
       Net Asn Ser Arg His Leu Thr Ile Thr Ile Ile Ala Gly Leu Ser Leu 1 5 10 15
35
       Phe Val Leu Thr Leu Gly Gly Cys Ser Val Ala Gln Gln Asp Thr Gln 20 25 30
       Trp Thr Leu Gly Gly Lys Leu Phe Thr Ser Ala Trp Ile Gln Arg Ser 35 40 45
40
       Ala Glu Tyr Gln Ala Leu Cys Ile Gln Ala Tyr Asn Ile Ala Thr Glu
50 60
       Arg Val Asp Ala Leu Pro Ala Glu Arg Lys Gln Gly Asp Arg Pro Tyr 65 70 70 75
       Ala Ile Val Thr Asp Ile Asp Glu Thr Ile Leu Asp Asn Thr Pro Asn
85 90 95
45
       Ser Val Tyr Gln Ala Leu Arg Gly Lys Asp Tyr Asp Glu Glu Thr Trp
100 105 110
       Gly Lys Trp Cys Ala Gln Ala Asp Ala Asp Thr Leu Ala Gly Ala Leu
115 120 125
50
       Ser Phe Phe Leu His Ala Ala Asn Lys Gly Ile Glu Val Phe Tyr Val
130 135 140
       Thr Asn Arg Arg Asp Asn Leu Arg Glu Ala Thr Leu Gln Asn Leu Gln 145 150 160

Arg Tyr Gly Phe Pro Phe Ala Asp Glu Glu His Leu Leu Thr Thr His 165 170 175
55
       Gly Pro Ser Asp Lys Glu Pro Arg Arg Leu Lys Ile Gln Glu Gln Tyr
180 185 190
       Glu Ile Val Leu Leu Ile Gly Asp Asn Leu Gly Asp Phe His His Fhe
195 200 205
60
       Phe Asn Thr Lys Glu Glu Ser Gly Arg Lys Gln Ala Leu Gly Leu Thr 210 215 220
       Ala Gly Glu Phe Gly Arg His Phe Ile Het Leu Pro Asn Pro Asn Tyr 225 230 240
65
       Gly Ser Trp Glu Pro Ala Trp Tyr Gly Gly Lys Tyr Pro Pro Leu Pro
245 250 255
       Glu Arg Asp Lys Ala Leu Lys Gln Leu His Ser Gln Asn Ser Arg
70
       (2) INFORMATION FOR SEQ ID NO:412
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 417 amino acids
                   (B) TYPE: amino acid
75
                   (D) TOPOLOGY: linear
```

```
(ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
 5
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
10
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...417
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:412
      lle Gl; Ser Ser Fro Leu Met Glu His Ala Ile Arg Val Ala Ala Gln
20 25 30
      Val Ala Pro Thr Asp Het Ser Val Leu Val Thr Gly Glu Ser Gly Ser 35 40 45
20
       Gly Lys Glu Phe Phe Pro Gln Ile Ile His Tyr Tyr Ser Ala Arg Lys
50 55 60
       His His Ser Tyr Ile Ala Val Asn Cys Gly Ala Ile Pro Glu Gly Thr 65 70 75 80
      Ile Asp Ser Glu Leu Phe Gly His Arg Lys Gly Ser Phe Thr Gly Ala 85 90 95
25
       Val Ser Asp Arg Lys Gly Tyr Phe Glu Glu Ala Ser Gly Gly Thr Ile
100 105 110
       Phe Leu Asp Glu Val Gly Glu Leu Pro Leu Pro Thr Gln Ala Arg Leu 115 125
30
       Leu Arg Val Leu Glu Thr Gly Glu Phe Ile Pro Val Gly Ala Ser Gln 130 135 140
       Ser Gln Lys Thr Asp Val Arg Ile Val Ala Ala Thr Asn Val Asn Leu
145 150 160
35
       Lys Glu Ala Val Ala Asn Gly Lys Phe Arg Glu Asp Leu Phe Phe Arg
165 170 175
       Leu Asn Thr Val Pro Ile Glu Val Pro Ala Leu Arg Met Arg Pro Asp
180 185 190
      Asp Val Pro Leu Phe Arg Arg Phe Ala Ala Asp Ser Ala Glu Lys
195 200 205
40
       Tyr Arg Net Pro Pro Leu Arg Leu Ser Asp Glu Ala Arg Thr Ile Leu
210 215 220
      Het Arg Tyr Arg Trp Pro Gly Asn Val Arg Glu Leu Arg Asn Ile Thr 225 230 240
      Asp Arg Leu Ser Ile Leu Glu Glu Glu Arg Thr Val Ser Ala Glu Thr
245 250 255
45
       Ile Thr Arg Tyr Leu Asp Ala Glu Gly Het Gln Asp Leu His Pro Val 260 265 270
       Val Ile Arg Arg Asn Glu Thr Thr Glu Ala Asp Lys Gln Ile Pro His
275 280 285
50
       Tyr Glu Arg Glu Ile Ile Tyr Gln Val Leu Tyr Asp Het Lys Lys Glu
290 295 300
       Ile Ala Asp Leu Lys Gly Met Het Asn Arg Leu Ala His His Glu Gln 305 310 315 320
55
       Pro Ser Trp Pro Val Gly Ser Asp Val Trp Gly Asn Asp Asp Lys Arg
325 330 335
       Thr Ala Asp Pro Lys Trp Gly Val Ser Thr His Lys Ala Pro Ile Ala 340 345 350
      Asn Ala Ala Glu Pro Val Glu Pro Ile Gln Glu Ala Ser Glu Tyr Thr
355 360 365
60
       Glu Asp Pro Val Ser Leu Glu Glu Val Glu Lys Lys Met Ile Ser Leu
370 375 380
      Ala Leu Glu Arg His Gly Gly Arg Arg Lys Gln Thr Ala Glu Glu Leu
385 390 395 400
      Lys Ile Ser Glu Arg Thr Leu Tyr Arg Lys Ile Lys Glu Tyr Gly Leu
405 410 415
65
70
       (2) INFORMATION FOR SEQ ID NO:413
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 602 amino acids
                  (B) TYPE: amino acid
75
                  (D) TOPOLOGY: linear
```

```
(ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
  5
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
10
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...602
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:413
15
       Het Ile Leu Leu Phe Gly Gly Thr Thr Glu Gly Arg Ala Ala Arg
       Val Leu Asp Glu Ala Gly Ser Pro Phe Phe Tyr Ser Thr Lys Gly Asn
20 25 30
       Leu Gl<br/>n Glu Ile Gl<br/>n Ser Ser His Gly His Arg Leu Thr Gly Ala Net 35 \phantom{\bigg|}40\phantom{\bigg|}
20
       Thr Val Ala Asp Het Val Ser Phe Cys Arg Lys Glu Glu Ile Arg Leu 50 60
       Ile Val Asp Ala Ala His Pro Phe Ala Glu Glu Leu His Ala Ser Val
65 70 75
25
       Ala Glu Als Ser Glu Gln Thr Gly Ile Pro Val Val Arg Tyr Glu Arg
85 90 95
       Gln Tyr Pro Pro Arg Glu Glu Gly Ile Val Trp Cys Ala Asn Tyr Asp
100 105 110
       Thr Ala Ala Glu Arg Met Leu Gly Asp Gly Val Gln Arg Leu Leu Met
115 120 125
30
       Leu Thr Gly Val Asn Thr Ile Pro Lys Leu Ala Ala Phe Trp Lys Glu
130 135 140
       Arg Thr Thr Phe Cys Arg Ile Leu Lys Arg Asp Glu Ser Val Ala Leu
145 150 155 160
35
       Ala Glu Lys Asn Gly Phe Pro Ala Glu Arg Ile Val Phe Phe Glu Pro
165 170 175
       His Ala Asp Glu Glu Leu Met Gln Ala Val Arg Pro Asp Ala Ile Ile
180 185 190
       Thr Lys Glu Ser Gly Glu Ser Gly Tyr Phe Arg Glu Lys Ile Glu Ala
195 200 205
40
       Ala Arg Arg Het Gly Ile Arg Ile Tyr Ala Val Val Arg Pro Pro Leu
210 215 220
       Pro Pro Ser Phe Ile Pro Val Gly Gly Pro Val Gly Leu Arg Arg Ala
225 230 235 240
45
       Val Glu Arg Leu Val Pro Gly Phe Phe Ser Leu Arg Ser Gly Phe Thr
245 250 250
       Thr Gly Thr Thr Ala Thr Ala Ala Val Val Ala Ala Het Tyr Arg Leu
260 265 270
       Het Gly Leu Gly Ser Leu Ala Glu Ala Pro Val Glu Leu Pro Ser Gly 275 280 285
50
       Glu Ile Val Ser Leu Pro Ile Ala Glu Ile Arg Glu Glu Glu Asp Ala
290 295 300
       Val Val Ser Ala Val Leu Lys Asp Ala Gly Asp Asp Pro Asp Val Thr
305 310 315 320
55
       Asn Gly Met Ala Val Cys Ala Thr Ile Arg Leu Asn Pro Glu His Glu
325 330 335
       Glu Val Arg Phe Leu Gln Gly Glu Gly Val Gly Val Val Thr Leu Pro
340 345 350
       Gly Leu Gl; Leu Glu Val Gly Gly Pro Ala Ile Asn Leu Val Pro Arg
355 360 365
60
       Arg Het Het Thr Ala Glu Val Arg Arg Leu Tyr Ala Gln Gly Gly Val 370 380
       Asp Ile Thr Ile Ser Val Pro Glu Gly Arg Glu Ala Ala Thr Gln Thr 385 390 395 400
65
       Phe Asn Pro Arg Leu Gly Ile Arg Asp Gly Ile Ser Ile Ile Gly Thr
405 415
       Ser Gl; Val Val Lys Pro Phe Ser Ala Glu Ala Phe Val Gly Ala Ile
420 425 430
       Arg Lys Gln Val Gly Ile Ala Thr Ala Leu Gly Ala Asn His Ile Val 435 440 445
70
       Leu Asn Ser Gly Ala Lys Ser Glu Arg Tyr Val Lys Gly Ala Tyr Pro
450 455 460
       Ala Leu Ile Pro Gln Ala Phe Val Gln Tyr Gly Asn Phe Val Gly Glu
465 470 475 480
75
       Ser Leu Ser Cys Val Ala Ser Phe Pro Ser Val Arg Ser Val Thr Val
```

```
Gly Ile Het Leu Gly Lys Ala Val Lys Leu Ala Glu Gly Tyr Leu Asp
500 505 510
       Thr His Ser Lys Lys Val Val Het Asn Arg Asp Phe Leu His Glu Leu
515 520 525
 5
       Ala Arg Gln Ala Gly Cys Ser Glu Asp Ile His Ala Ile Ile Asp Ser
530 540
       Leu Asn Leu Ala Arg Glu Leu Trp Thr Met Pro Ser Ala Glu Asp Ser 545 550 560
10
       Asp Arg Leu Leu Arg Lys Ile Ala Glu Arg Ser Trp Glu Thr Cys Arg 565 570 575
       Pro Ser Val Pro Ser Ala Glu Leu Glu Leu Leu Leu Ile Asp Glu Ser 580 585 590
       Gly Ala Ile Arg Phe Arg Ile Gly Gly Glu
15
       (2) INFORMATION FOR SEQ ID NO:414
             (i) SEQUENCE CHARACTERISTICS:
20
                   (A) LENGTH: 443 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
25
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
30
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...443
35
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:414
       Met Leu Arg Thr Phe Arg Ile Gly Gly Ile His Pro Pro Glu Asn Lys
                                               10
       Leu Ser Ala Gly Lys Pro Val Glu Val Leu Pro Ile Pro Ser Gln Val
20 25 30
40
       Val Ile Pro Leu Gly Gln His Ile Gly Ala Pro Ala Thr Ala Thr Val
35 40
       Lys Lys Gly Asp Glu Val Lys Val Gly Thr Ile Ile Ala Gln Ala Gly 50 55
       Gly Phe Val Ser Ala Asn Ile His Ser Ser Val Ser Gly Lys Val Leu
65 70 80
45
       Lys Ile Asp Asn Val Tyr Asp Ser Ser Gly Tyr Pro Lys Pro Ala Val
85 90
       Phe Ile Ser Val Glu Gly Asp Glu Trp Glu Glu Gly Ile Asp Arg Ser 100 105 110
50
       Pro Ala Ile Val Lys Glu Cys Asn Leu Asp Ala Lys Glu Ile Val Ala
115 120 125
       Lys Ile Ser Ala Ala Gly Ile Val Gly Leu Gly Gly Ala Thr Phe Pro
130 135 140
55
       Thr His Val Lys Leu Ser Pro Pro Pro Gly Asn Lys Ala Glu Ile Leu
145 150 150
       The The Asn Ala Val Glu Cys Glu Pro Tyr Leu Thr Ser Asp His Val
165 170 175
       Leu Met Leu Glu His Gly Glu Glu Ile Het Ile Gly Val Ser Ile Leu
180 185 190
Met Lys Ala Ile Gly Val Ser Ile Leu
60
      Met Lys Ala Ile Gln Val Asn Lys Ala Val Ile Gly Val Glu Asn Asn 195 200 205
       Lys Lys Asp Ala Ile Ala His Leu Thr Lys Leu Ala Thr Ala Tyr Pro210 \\ 215 \\ 220
65
       Gly Ile Glu Val Het Pro Leu Lys Val Gln Tyr Pro Gln Gly Gly Glu
225 230 235 240
       Lys Gln Leu Ile Asp Ala Val Ile Arg Lys Gln Val Lys Ser Gly Ala
245 250 255
      Leu Pro Ile Ser Thr Gly Ala Val Val Gln Asn Val Gly Thr Val Phe 260 265 270
70
      Ala Val Tyr Glu Ala Val Gln Lys Asn Lys Pro Leu Val Glu Arg Ile
275 280 285
       Val Thr Val Thr Gly Lys Lys Leu Ser Arg Pro Ser Asn Leu Leu Val
290 295 300
75
       Arg Ile Gly Thr Pro Ile Ala Ala Leu Ile Glu Ala Ala Gly Gly Leu
```

```
Pro Glu Asn Thr Gly Lys Ile Ile Gly Gly Gly Pro Met Met Gly Arg
325 330 335
        Ala Leu Leu Ser Pro Asp Val Pro Val Thr Lys Gly Ser Ser Gly Val 340 345 350
 5
        Leu Ile Leu Asp Arg Glu Glu Ala Val Arg Lys Pro Het Arg Asp Cys 355 360 365
        Ile Arg Cys Ala Lys Cys Val Gly Val Cys Pro Het Gly Leu Asn Pro
370 380
10
        Ala Phe Leu Met Arg Asp Thr Leu Tyr Lys Ser Trp Glu Thr Ala Glu
385 390 395 400
        Lys Gly Asn Val Val Asp Cys Ile Glu Cys Gly Ser Cys Ser Phe Thr
405 415
        Cys Pro Ala Asn Arg Pro Leu Leu Asp Tyr Ile Arg Gln Ala Lys Lys
420 425 430
15
        Thr Val Het Gly Ile Gln Arg Ala Arg Lys Gln
435
        (2) INFORMATION FOR SEQ ID NO:415
20
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 479 amino acids
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
25
            (ii) HOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KET: misc_feature
35
                   (B) LOCATION 1...479
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:415
       Met Lys Arg Ile Gln Leu Thr Leu Ile Ala Leu Phe Ala Ala Val Ala 1 \phantom{a} 10 \phantom{a} 15
40
       Gly Leu Val Ala Gln Asn Ala Tyr Glu Gly Val Ile Ser Tyr Lys Ile
20 25 30
       Ser Leu Asp Lys Thr Gly Asn Lys Val Val Leu Asn Gly Ala Ala Asp 35 40 45
45
       Net Ser Asn Leu Lys Leu Lys Ser Thr Gln Net Ile Ile Val Thr Pro
50 55 60
       Ile Leu Arg Ser Glu Asp Gly Thr Ser Arg Val Glu Phe Pro Ser Val 65 \phantom{000}75\phantom{000} 70 \phantom{0000}80\phantom{000}
       Val Ile Thr Gly Arg Asn Arg Thr Lys Ala Leu Lys Arg Glu Ile Ala
85 90 95
50
       Phe Ser Ser Ala Leu Pro Gln Ala Lys His Ala Ala Gln Tyr Ile Arg
100 105 110
       Arg His Asn Gly Lys Ser Glu Gln Phe Ala Phe Thr Gly Glu His Ala
115 120 125
55
       Tyr Ala Ser Trp Met Het Asp Ala Lys Phe Val Val Arg Glu Glu Val
130 135 140
       Arg Gly Cys Ala Lys Cys Pro Val Gly Leu Ser Ser Asn Ile Val Pro 145 150 160
       Phe Asp Pro Leu Phe Asn Pro Ala Glu Ala Pro Tyr Leu Leu Ala His
165 170 175
60
       Ile Thr Pro Ala Glu Glu Val Glu Lys Gln Arg Glu Ser Ser Phe Asp
180 185 190
       Ala Tyr Ile Asn Phe Lys Val Asn Lys Ala Asp Val Leu Pro Glu Tyr 195 200
65
       Arg Asn Asn Lys Ala Glu Leu Glu Lys Ile Lys Glu Phe Val Ser Thr 210 215 220
       Val Lys Ala Asn Pro Asn Tyr Ser Val Asn Lys Met Ile Ile Glu Gly 235 230 235
       Phe Ala Ser Pro Glu Ala Ser Ile Ala His Asn Lys Ala Leu Ser Glu
245 250 255
70
       Arg Arg Ala Lys Arg Leu Ala Glu Glu Leu Val Arg Lys Tyr Gly Lys 260 265 270
       Thr Leu Pro Asn Ile Thr Thr Glu Phe Gly Gly Glu Asp Trp Lys Gly 275 280 285
75
       Leu Lys Leu Ala Ile Glu Lys Ser Asp Ile Ala Asp Arg Asp Arg Val
```

```
295
                                                          300
       Leu Glu Ile Ile Asn Ser Asp Lys Tyr Ala Asp Asp Asp Ala Arg Glu
305 310 315 320
       Gln Ala Leu Lys Gln Leu Ser Ser Tyr Arg Tyr Ile Leu Asp Gln Ile
325 330 335
       Tyr Pro Asn Leu Arg Arg Asn Thr Ile Thr Met Gly Tyr Ile Val Arg 340 345 350
       Asp Tyr Thr Leu Glu Glu Ala Arg Glu Ile Ile Lys Thr Ala Pro Lys
355 360 365
10
      Glu Leu Ser Glu Ala Glu Het Tyr Arg Val Ala Het Ser Tyr Pro Glu 370 375 380
       Gly His Gln Glu Arg Leu Phe Ala Leu Asn Thr Thr Leu Lys Tyr Phe
385 390 395 400
       Pro Glu Ser Val Thr Gly Arg Ile Asn Leu Ala Val Ala Ala Phe Asn
15
                        405
                                              410
       Gly Gly Asp Val Gln Gln Ala Ile Ala Leu Leu Ser Pro Ile Gln Thr
420 425 430
       Glu Lys Gly Val Ser Asm Ile Leu Gly Ala Ala Tyr Ala Arg Thr Gly
435 446 445
20
       Asp Phe Ala Arg Ala Glu Thr Phe Phe Arg Lys Ala Val Ala Glu Gly
450 455 460
                            455
       Asp Ala Asn Ala Gln Arg Asn Leu Asp Met Leu Leu Gly Lys Lys
25
       (2) INFORNATION FOR SEQ ID NO:416
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 383 amino acids
                   (B) TYPE: amino acid
30
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
35
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
40
                   (A) HAME/KEY: misc_feature
                   (B) LOCATION 1...383
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:416
45
       Met Ala Glu Lys Arg Asp Tyr Tyr Glu Val Leu Gly Val Ser Lys Asn
                                               10
       Ala Thr Asp Asp Glu Leu Lys Lys Ala Tyr Arg Lys Lys Ala Ile Glu 20 $25$
       Tyr His Pro Asp Lys Asn Pro Gly Asp Lys Glu Ala Glu Glu His Phe 35 \hspace{1cm} 40 \hspace{1cm} 45
50
       Lys Glu Val Ala Glu Ala Tyr Asp Val Leu Ser Asp Pro Gln Lys Arg 50 60
       Ser Gln Tyr Asp Gln Phe Gly His Ala Gly Leu Gly Gly Ala Ala Gly 65 70 75
       Gly Gly Phe Ser Gly Gly Het Ser Het Glu Asp Ile Phe Ser Arg
85 90 95
55
       Phe Gly Asp Leu Phe Gly Gly Phe Gly Gly Phe Gly Gly Phe Ser Asp 100 \hspace{1cm} 105 \hspace{1cm} 110
       Net Gly Gly Ser Arg Arg Arg Val Arg Arg Gly Ser Asp Leu Arg
115 120
60
       Val Arg Val Lys Leu Ser Leu Ala Asp Ile Ser Lys Gly Val Glu Lys 130 135 140

Lys Val Lys Val Lys Gln Val Val Cys Ser Lys Cys Arg Gly Asp 145 150 150 155 160
65
       Gly Thr Glu Glu Ala Asn Gly Lys Thr Thr Cys Gln Thr Cys His Gly
165 170 175
       Thr Gly Val Val Thr Arg Val Ser Asn Thr Phe Leu Gly Ala Met Gln 180 190
       Thr Gln Ser Thr Cys Pro Thr Cys His Gly Glu Gly Glu Ile Ile Thr
195 200 205
70
       Lys Pro Cys Ser Lys Cys Lys Gly Glu Gly Val Glu Ile Gly Glu Glu 210 215 220
       Val Ile Ser Phe His Ile Pro Ala Gly Val Ala Glu Gly Met Gln Met 225 230 230 235
                                                  235
75
       Ser Val Asn Gly Lys Gly Asn Ala Ala Pro Arg Gly Gly Val Asn Gly
```

```
Asp Leu Ile Val Val Ile Ala Glu Glu Pro Asp Pro Asn Leu Ile Arg
260 265 270
       Asn Gly Asn Asp Leu Ile Tyr Asn Leu Leu Ile Ser Val Pro Leu Ala
275 280 285
  5
       Ile Lys Gly Gly Ser Val Glu Val Pro Thr Ile Asp Gly Arg Ala Lys
290 295 300
       Ile Arg Ile Glu Ala Gly Thr Gln Pro Gly Lys Met Leu Arg Leu Arg
305 310 315 320
       Asn Lys Gl; Leu Pro Ser Val Asn Gl; Tyr Gl; Met Gl; Asp Gln Leu
325 330 335
 10
       Val Asn Val Asn Val Tyr Ile Pro Glu Ser Ile Asp Ala Lys Asp Glu
340 345 350
       Gln Ala Ile Ala Ala Met Glu Asn Ser Asp Ser Phe Lys Pro Thr Asp
355 360 365
15
       Ala Ala Arg Lys Asp Ile Asp Lys Lys Tyr Arg Glu Met Leu Asp
370 375 380
       (2) INFORMATION FOR SEQ ID NO:417
20
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 293 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
35
                  (B) LOCATION 1...293
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:417
       40
       Leu Ser Cys Ser Ser Asn Asn Lys Asp Leu Glu Asn Lys Gly Glu Ala 20 25 30
       Thr Leu Leu Val Thr Phe Gly Ser Ser Tyr Lys Ala Pro Arg Glu Thr
35 40 45
       Tyr Ala Lys Ile Glu Lys Thr Phe Ala Ala Ala Tyr Pro Asp Gln Arg
50 60
45
       Ile Ser Trp Thr Tyr Thr Ser Ser Ile Ile Arg Lys Lys Leu Ala Gln 65 70 80
       Gln Gly lie Tyr lle Asp Ala Pro Asp Glu Ala Leu Glu Lys Leu Ala
85 90
50
       Arg Leu Gly Tyr Lys Lys Ile Asn Val Gln Ser Leu His Val Ile Pro 100 \,
       Gly Arg Glu Tyr Asp Glu Met Ile Asp Phe Val Asn Lys Phe Lys Ala
115 120 125
55
       Ala His Ser Asp Ile Thr Val Lys Val Gly Ala Pro Leu Phe Asp Thr
130 135 140
      60
      Glu His Ala Ala Asn Asp Arg Tyr Ala Arg Ile Asn Lys Ile Met Lys
180
185
190
Asn Tyr Ser Lys Phe Met Ile Val Gly Thr Val Glu Ser Asp Pro Ser
195
200
205
65
      Ile Asn Asp Val Ile Ala Glu Leu Lys Glu Thr Gly Ala Thr Ala Val
210 215 220
      Thr Het Het Pro Leu Het Ser Val Ala Gly Asp His Ala Thr Asn Asp 225 230 235 240
      Het Ala Gly Asp Glu Asp Asp Ser Trp Lys Thr Leu Leu Thr Asn Ala 245 250 255
70
      Gly Tyr Thr Val Ser lie Asp Lys Leu Asp Asn Gly Asn Phe Ser Ala 260 265 270
      Leu Gly Asp Ile Glu Glu Ile Arg Asn Ile Trp Leu Lys His Het Lys
275 280 285
75
      Ala Thr Ser Ala Arg
```

```
290
```

```
(2) INFORMATION FOR SEQ ID NO:418
  5
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 356 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 10
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
 15
                  (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...356
 20
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:418
       Het Thr Ser Val Ser His Leu Arg Thr Ile Ser Val Ala Gly Ile Leu
25
       Ala Ala Leu Gly Gly Ala Val Leu Ile Leu Phe Gly Val Asn Leu Phe 20 25 30
       Leu Gly Ser Val Ala Ile Pro Net Ser Glu Ile Phe Arg His Leu Phe
                                     40
       Ser Asp Arg Pro Glu Gly Gly Glu Ala Leu Val His Tyr Asn Ile Leu
50 55 60
30
       Trp Lys Ser Arg Leu Pro Glu Ala Leu Thr Ala Ala Phe Ala Gly Ala
65 70 80
       Gly Leu Ser Val Ser Gly Leu Gln Het Gln Thr Val Phe Arg Asn Pro
85 90 95
35
       Leu Ala Gly Pro Ser Val Leu Gly Ile Ser Ser Gly Ala Ser Leu Gly 100 105 110
       Val Ala Leu Val Val Leu Leu Ser Gly Ser Leu Gly Gly Val Ala Leu
115 120 125
       Ser Ser Leu Gly Tyr Met Gly Glu Val Ala Met Asn Ile Ala Ala Ala 130 135 140
40
       Val Gly Ser Leu Ala Val Met Gly Leu Ile Val Phe Val Ser Thr Lys
145 150 155 160
       Val Arg Ser His Val Thr Leu Leu Ile Gly Val Met Ile Gly Tyr
165 170 175
45
       Val Ala Thr Ala Val Ile Gly Val Phe Lys Phe Phe Ser Ile Glu Glu
180 185 190
       Asp Ile Arg Ala Tyr Val Ile Trp Gly Leu Gly Ser Phe Ser Arg Ala
195 200 205
       Thr Asp Ser Gln Leu Ser Phe Phe Ala Ile Leu Het Leu Ile Phe Ile
210 215 220
50
       Pro Ala Gly Het Leu Leu Val Lys Gln Leu Asn Leu Leu Leu Gly 235 230 235
       Glu Ser Tyr Ala Arg Asn Leu Gly Leu Asn Thr Arg Arg Ala Arg Leu
245 250 255
55
       Leu Val Ile Ser Ser Ala Gly Leu Leu Ile Ala Thr Val Thr Ala Tyr
260 265 270
       Cys Gly Pro Ile Gly Phe Leu Gly Met Ala Val Pro His Leu Ala Arg
275 280 285
       Val Ile Phe His Thr Ser Asp His Arg Ile Leu Met Pro Ala Thr Cys
290 295 300
60
       Leu Ile Gl; Ser Ala Leu Ala Leu Phe Cys Asn Ile Ile Ala Arg Met
305 310 315 320
       Pro Gly Phe Glu Gly Ala Leu Pro Val Asn Ser Val Thr Ala Leu Val
325 330 335
65
       Gly Ala Pro Ile Ile Val Thr Val Leu Phe Arg Arg Arg Phe Lys
       Glu Glu Thr Asp
               355
70
       (2) INFORMATION FOR SEQ ID NO:419
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 757 amino acids
                  (B) TYPE: amino acid
75
                  (D) TOPOLOGY: linear
```

```
(ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
10
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...757
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:419
15
       Met Arg Thr Lys Thr Ile Phe Phe Ala Ile Ile Ser Phe Ile Ala Leu
       Leu Ser Ser Ser Leu Ser Ala Gln Ser Lys Ala Val Leu Thr Gly Ser
20 25 30
       Val Ser Asp Ala Glu Thr Gly Glu Pro Leu Ala Gly Ala Arg Ile Glu 35 40 45
20
       Val Lys His Thr Asn Ile Val Ala Gly Ala Asp Ala Gly Gly His Phe 50 \hspace{1cm} 55
       Glu Île Lys Asn Leu Pro Ala Gly Gln His Thr Île Île Cys Ser Leu 65 70 80
25
       Gly Gly Tyr Gly Gln Lys Glu Glu Val Val Ala Ile Glu Ala Gly Gln
85 90 95
       Thr Lys Thr Ile Ser Phe Ala Leu Arg Leu Arg Thr Asn Asn Leu Glu 100 105 110
       Glu Val Val Thr Gly Thr Gly Thr Arg Tyr Arg Leu Val Asp Ala
115 120 125
30
       Pro Val Ala Thr Glu Val Leu Thr Ala Lys Asp Ile Ala Ser Phe Ser
130 135 140
       Ala Pro Thr Ser Glu Ala Leu Leu Gln Gly Leu Ser Pro Ser Phe Asp
145 150 155 160
35
       Phe Gly Pro Asn Leu Met Gly Ser Phe Het Gln Leu Asn Gly Leu Ser
165 170 175
       Ser Lys Tyr Ile Leu Ile Leu Ile Asp Gly Lys Arg Val Tyr Gly Asp
180 185 190
       Val Gly Gly Gln Ala Asp Leu Ser Arg Ile Ser Pro Asp Gln Ile Glu
195 200 205
40
       Arg Ile Glu Leu Val Lys Gly Ala Ser Ser Ser Leu Tyr Gly Ser Asp
210 215 220
       Ala Ile Ala Gly Val Ile Asn Val Ile Thr Lys Lys Asn Thr Asn Arg
225 230 230 235 240
       Leu Ser Ala Tyr Thr Ser His Arg Ile Ser Lys Tyr Asn Asp Arg Gln 245 250
       Thr Asn Thr Ser Leu Asp Ile Asn Ile Gly Lys Phe Ser Ser Asn Thr 260 265 270
       Ash Tyr Phe Phe Tyr His Thr Asp Gly Trp Gln Ash Ser Pro Phe Glu
275 280 285
50
       Ile Lys Lys Lys Gly Ser Gly Glu Pro Val Leu Glu Glu Thr Tyr 290 295 300
       Lys Lys Thr Phe Arg Ala Gln Glu Asn Gln Gly Val Ser Gln Ser Leu
305 310 315 320
       Ser Tyr Tyr Ala Thr Asn Asn Leu Ser Phe Ser Gly Asn Val Gln Tyr 325 330 335
55
       Asn Lys Arg Gln Ile Phe Thr Pro Thr Phe Ser Glu Lys Lys Ala Tyr
340 345 350
       Asp Het Asp Tyr Arg Ala Leu Thr Ala Ser Leu Gly Thr Asn Tyr Leu 355 360 365
60
       Phe Pro Asn Gly Leu His Thr Leu Ser Phe Asp Ala Val Tyr Asp Arg 370 380
       Phe Arg Phe Gly Tyr Leu Tyr His Asp Lys Asp Ser Ser Glu Ser Leu 385 390 395 400
65
       Ile Asn Asn Gln Gly Gln Thr Glu Gln Pro Thr Phe Pro Gly Gln 405 410 415
       Leu Arg Asn Lys Asn Asp Gln Ile Arg Tyr Thr Ala Glu Ala Arg Gly 420 425 430
       Val Phe Thr Leu Pro Tyr Ala Gln Lys Leu Thr Gly Gly Leu Glu Tyr
435 440 445
70
       Phe Arg Glu Glu Leu Ile Ser Pro Tyr Asn Leu Ile Thr Asp Lys Ala
450 460
                                455
       Asp Ala Ser Thr Leu Ser Ala Tyr Val Gln Asp Glu Trp Lys Pro Leu
465 470 475 480
```

Asp Trp Phe Asn Net Thr Ala Gly Phe Arg Leu Val His His Gln Glu

75

| | Phe | Gly | Thr | | 485 Net | Thr | Pro | Lys | | 490 Ser | Ile | Leu | Ala | | 495 Tyr | Gly |
|------------|------------------------|------------|------------|------------------------|------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 5 | Pro | Leu | Asn 515 | 500 Phe | Arg | Ala | Thr | Tyr 520 | 505 Ala | Asn | Gly | T;·r | Lys 525 | 510 Thr | Pro | Thr |
| Ü | Leu | Lys 530 | | Leu | Phe | Ala | Arg 535 | | Glu | Leu | Thr | Thr 540 | | Gly | Ser | His |
| | Asn 545 | Leu | T).r | Leu | Gly | Asn 550 | Ala | Asp | Leu | Lys | Pro 555 | Gln | Net | Ser | Агр | Tyr 560 |
| 10 | - | | | | Leu 565 | | | | | 570 | | | | | 575 | |
| | | | | 580 | Asn | | | | 585 | | | | | 590 | | |
| 15 | | | 595 | | Glu | | | 600 | | | | | 605 | | | |
| | _ | 610 | | | Gly. | _ | 615 | _ | | _ | | 620 | | | | |
| | 625 | Ala | ser | 116 | Gly | 630 | GIĀ | TIE | ьуѕ | Leu | 635 | Ala | GT7. | 1 Y L | ser | 640 |
| 20 | | Glu | Ala | Lys | Asn 645 | | Gln | Thr | Asp | Glu 650 | | Leu | Glu | Gly | Ala 655 | |
| | Arg | His | Arg | Ala 660 | Asn | Val | His | Ala | Asp 665 | Trp | Val | His | T7.r | Trp 670 | Gly | Gln |
| 25 | Tyr | Arg | Leu 675 | Gly | Val | Ser | Leu | Phe 680 | Gly | Arg | Ile | Gln | Ser 685 | Glu | Arg | T/r |
| | $T\gamma r$ | | Asp | Glγ | Asn | Ala | | A.s.p | Tyr | Thr | Leu | | Arg | Leu | Ala | Thr |
| | | 690 His | Arg | Phe | Ala | | 695 Phe | Arg | His | Ile | Ile 715 | 700 Leu | Asp | Gly | Thr | |
| 30 | 705 Gl _y | IJ€ | Asp | Asn | Leu 725 | 710 Phe | Asp | туг | Val | Asp 730 | | Arg | Pro | Met | Gly 735 | 720 Val |
| | Asn | Tyr | Ala | Thr 740 | Val | Thr | Pro | Gly | Arg 745 | | Phe | Phe | Ala | Gln 750 | | Ala |
| 35 | Ile | Arg | Phe 755 | Asn | Asn | | | | | | | | | | | |
| | (2) | INFO | ORI1AT | rion | FOR | SEQ | IDI | 10:42 | 20 | | | | | | | |
| | | / - | 000 | OLUENIA | יים כיו | א מ מ ע ו | ומשיים | COTA | ٠ | | | | | | | |
| 4 0 | | 11 | () | 4) L | CE CH ENGTH YPE: | H: 33 | 31 ar | nino | | is | | | | | | |
| | | | | | OPOLO | | | | | | | | | | | |
| 45 | | (ii) |) MOI | LECUI | LE TY | PE: | prot | cein | | | | | | | | |
| | | (iii |) HTI | POTHI | ETICA | AL: Y | YES | | | | | | | | | |
| | | (vi | | | AL SC RGANI | | | obwre | omon: | ഭവ | ingi | zalis | = | | | |
| 50 | | | ٠. | ., | | | 2021 | , | | 9- | | | - | | | |
| | | (ix) | | ATURI A) NA | E: 414E/1 | ŒY: | misc | e fea | ature | = | | | | | | |
| | | | (E | 3) L(| OCAT] | 101 | 13 | 331 | | | | | | | | |
| 55 | | (xi |) SE(| QUENG | CE DE | ESCR: | IPTIC |)H: : | SEQ : | ID HO | 0:420 |) | | | | |
| | Net 1 | Thr | Asp | Asn | Lys 5 | Gln | Arg | Asn | Ile | Val 10 | Phe | Pro | Ala | Phe | Leu 15 | Leu |
| 60 | Leu | Leu | G17. | Val 20 | Ile | Ala | Val | Уаl | Thr 25 | | Val | Gly | Phe | Phe 30 | | Leu |
| | Arg | Pro | Ala 35 | Glu | Glu | Ile | Ile | Gln 40 | Gl y | Gln | Ile | Glu | Val 45 | Thr | Glu | туг |
| 0.5 | Arg | Val 50 | Ser | Ser | Lys | Val | Pro 55 | Gly | Arg | Ile | Γλε | Glu 60 | Leu | Arg | Val | Ser |
| 65 | 65 | | | | Val | 70 | | _ | - | | 75 | | | | | 80 |
| | | | | | Ala 85 | | | | | 90 | - | | | | 95 | |
| 70 | | | | 100 | Asn | | _ | | 105 | _ | | | - | 110 | | |
| | | | 115 | | Tyr | | | 120 | | | | | 125 | | | |
| | TIG | wra | ınr | $r\lambda \varepsilon$ | Thr | HIS | | Arg | val | OTU | Asn | | Tyr | Asp | GID | PT ?. |
| 75 | | 130 | | | Gln | | 135 | | | | | 140 | | | | |

```
150
        Ala Ile Ala Thr Gln Lys Ala Ala Glu Ala Gln Tyr Asn Met Ala Arg
165 170 175
        Asn Gly Ala Glu Arg Glu Asp Lys Leu Ala Ala Ser Ala Leu Val Asp
180 185 190
 5
        Arg Ala Arg Gly Ala Val Ala Glu Val Glu Ser Tyr Ile Asn Glu Thr
195 200 205
        Tyr Leu Ile Ala Pro Arg Ala Gly Glu Val Ser Glu Ile Phe Pro Lys
210 215 220
        Ala Gly Glu Leu Val Gly Thr Gly Ala Pro Ile Het Asn Ile Ala Glu
225 230 240
10
        Het Gly Asp Het Trp Ala Ser Phe Ala Val Arg Glu Asp Phe Leu Ser
245 250 255
        Ser Net Thr Net Gly Ala Val Leu Glu Thr Val Val Pro Ala Leu Asn
260 265 270
15
        Glu Glu Lys Val Arg Phe Lys Ile Thr Phe Ile Lys Asn Met Gly Thr
275 280 285
       Tyr Ala Ala Trp Lys Ala Thr Lys Thr Thr Gly Gln Tyr Asp Leu Lys 290 295 300

Thr Phe Glu Val Lys Ala Thr Leu Ala Asp Lys Asp Lys Ala Gln Lys 305 310
20
        Leu Arg Pro Gly Met Ser Val Ile Ile Arg Lys
25
        (2) INFORMATION FOR SEQ ID NO:421
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 267 amino acids
                     (P) TYPE: amino acid
30
                     (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
35
             (vi) ORIGINAL SOURCE:
                     (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
40
                    (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...267
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 421
45
       Net Arg Ile Val Ser Asn Phe Leu Phe Val Ser Phe Ser Val Leu Leu 1 5 10 15
        Phe Ala Ser Cys Arg Ser Gln Arg Glu Lys Val Val Tyr Leu Gln Asp
20 25 30
        Ile Gln Thr Phe Asn Arg Glu Ile Ile Ala Lys Pro Tyr Asp Val Lys
35 40 45
50
        Ile Glu Lys Asp Asp Val Leu Asn Ile Leu Val Ser Ser Arg Asp Pro
50 60
        Glu Leu Ser Thr Pro Tyr Asn Gln Val Leu Thr Thr Arg Ala Leu Ala
65 70 80
55
        Arg Asn Gly Tyr Gly Thr Asn Ser Asn Glu Gly Phe Leu Val Asp Ser
85 90 95
        Lys Gly Tyr Ile Asn Tyr Pro Ile Leu Gly Gln Ile Tyr Val Glu Gly
100 105
       Leu Thr Arg Thr Glu Leu Glu Lys Glu Ile Gln Lys Arg Ile Ile Ser 115
Ser Gly Phe Ile Lys Asp Pro Thr Val Thr Val Gln Leu Gln Asn Phe 130
Lys Val Ser Val Leu Gly Glu Val Asn His Pro Gly Ser Het Ser Val 145

150
155
169
60
        Lys Gly Glu Arg Ile Thr Leu Leu Glu Ala Ile Gly Het Ala Gly Asp
165 175
65
       Leu Thr Ile Tyr Gly Arg Arg Asp Arg Val Phe Val Ile Arg Glu Thr 180 \hspace{1cm} 185 \hspace{1cm} 190 \hspace{1cm}
       Asp Gly His Arg Glu Val Phe Gln Thr Asp Leu Arg Lys Ala Asp Leu
195 200 205
70
        Leu Ala Ser Fro Val Tyr Tyr Leu His Gln Asn Asp Val Ile Tyr Val
210 215 220
        Glu Pro Asn Asp Lys Lys Thr Gln Met Ser Glu Ile Asn Gin Asn Asn
225 230 240
75
        Asn Val Asn Val Trp Leu Ser Val Thr Ser Thr Leu Val Ser Ile Ser
```

```
245
                                                                                                250
                                                                                                                                              255
              Thr Leu Thr Ile Thr Ile Ile Asp Lys Thr Lys
   5
               (2) INFORMATION FOR SEQ ID NO: 422
                          (i) SEQUENCE CHARACTERISTICS:
                                      (A) LENGTH: 569 amino acids
                                      (B) TYPE: amino acid
10
                                      (D) TOPOLOGY: linear
                        (ii) MOLECULE TYPE: protein
                     (iii) HYPOTHETICAL: YES
15
                        (vi) ORIGINAL SOURCE:
                                     (A) ORGANISM: Porphyromonas gingivalis
                        (im) FEATURE:
20
                                     (A) NAME/KEY: misc_feature
                                     (B) LOCATION 1...569
                        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 422
25
              Het Lys Lys Thr Asn Leu Phe Leu Ser Leu Leu Val Ile Phe Ile Thr
              Gly Ser Phe Het Thr Ala Cys Ala Gln Lys Ser Lys Thr Asn Lys Leu 20 \hspace{1cm} 25 \hspace{1cm} 30
              Thr Glu Glu Asp Arg Ser Arg Asn Glu Tyr Val Gln Ser Met Asp Val 35 40 45
30
              Leu Ser Asn Ile Ile Gly Asn Val Arg Leu Tyr Phe Val Asp Thr Ile 50 \hspace{1cm} 55
              Ser Ile Lys His Met Thr Arg Arg Gly Ile Asp Ala Met Leu Gly Gly 65 70 75 80
35
              Leu Asp Pro Tyr Thr Glu Tyr Ile Pro Tyr Glu Glu Met Asp Glu Leu
85 90 95
              Lys Leu Net Thr Thr Gly Glu Tyr Ala Gly Val Gly Ala Ile Ile Ser
100 105 110
              Gln Arg Pro Asp Ser Ala Val Ile Ile Gln Arg Pro Met Glu Gly Met
115 120 125
40
              Pro Ala Asp Glu Ala Gly Leu Ile Ala Gly Asp Arg Ile Leu Thr Ile
130 135
              Asp Gly Lys Asp Phe Arg Lys Ser Thr Thr Pro Lys Val Ser Gln Ala
145 150 160
              Leu Lys Gly Ile Ala Gly Thr Val Ala Lys Val Thr Val Met Arg Tyr
165 170 175
Gly Glu Thr Lys Pro Arg Thr Phe Ser Val Lys Arg Gln Lys Val Ile
180 185 190
45
              Het Asn Ser Val Thr Tyr Ser Gly Met Leu Asp Gly Ser Ile Gly Tyr
195 200 205
50
              Ile Arg Leu Asn Asn Phe Thr Asp Lys Ser Ala Glu Glu Val Arg Thr
210 215 220
              Ala Leu Leu Asp Leu Arg Asp Lys Gln Gl\gamma Ala Lys Gly Leu Ile Leu 225 230 235 240
55
              Asp Leu Arg Gly Asn Gly Gly Gly Leu Met Gln Ala Ala Ile Glu Ile
245 250 255
              Val Asn Leu Phe Val Pro Lys Gly Lys Glu Val Val Thr Thr Lys Gly 260 265 270
              Arg Ile Ala Glu Ser Ala Ser Val Phe Arg Thr Leu Thr Glu Pro Ile
275 280 285
60
              Asp Thr Lys Leu Pro Ile Val Val Leu Ile Asp Gly Gln Ser Ala Ser
290 295 300
              Ser Ser Glu Ile Val Ala Gly Ala Leu Gln Asp Het Asp Arg Ala Val
305 310 315 320
              Leu Het Gly Gln Lys Ser Tyr Gly Lys Gly Leu Val Gln Thr Thr Arg
325
330
335
Gln Leu Pro Tyr Asn Gly Val He Lys Leu Thr Thr Ala Lys Tyr Tyr
340
345
350
65
              | The | Pro | Ser | Gly | Arg | Cys | The | Gln | Arg | Leu | Asp | Tyr | Ser | Arg | Thr | Asn | 355 | 360 | 365 | 365 | 365 | Arg | Thr | Gly | Het | Ala | Thr | Ala | The | Ala | Thr 
70
              Tyr Thr Ala Ala Gly Arg Arg Val Glu Asp Ala Gly Gly Ile Leu Pro
                                                        390
                                                                                                      395
75
              Asp Ile Glu Val Lys Gln Asp Thr Ala Ala Thr Leu Leu Tyr Tyr Met
```

```
405
                                               410
       Ala Ile Asn Asn Asp Val Phe Asp Phe Val Thr Gly Tyr Val Leu Lys
420 425
       His Lys Thr Ile Ala Lys Pro Giu Asp Phe Ser Ile Thr Asn Glu Asp 435 440 445
 5
       Tyr Ala Ala Phe Cys Lys Met Met Glu Glu Lys Lys Phe Asp Tyr Asp 450 450 460
       Arg Gln Ser Gly Lys Met Leu Asp Lys Leu Glu Glu Leu Ala Lys Ile
465 470 475 480
       Glu Gly Tyr Leu Pro Glu Ala Asn Ser Glu Leu Lys Ala Leu Arg Glu 485 490 495
10
       Lys Leu Lys Pro Asn Leu Ser Arg Asp Leu Leu Arg Phe Lys Lys Glu 500 505 510
       Ile Thr Asn Tyr Leu Asn Asn Glu Ile Val Thr Arg Tyr Tyr Glu
515 520 525
15
       Arg Gly Ser Ile Arg Gln Ser Leu Pro Glu Asp Lys Val Val Lys Glu
530 535 540
       Ala Ile Lys Leu Leu Lys Asp His Pro Glu Gln Ile Arg Gln Ile Leu
545 550 560
20
       Ala Ala Pro Lys Ala Glu Asn Lys Gly
       (2) INFORMATION FOR SEQ ID NO:423
25
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 981 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
30
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
35
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...981
40
           (::i) SEQUENCE DESCRIPTION: SEQ ID NO:423
       Het Gln Asn Lys Gly Phe Val Ile Val Ile Thr Ser Ala Leu Ala Ile 1 5 10 15
45
       Ile Cys Ala Phe Tyr Leu Ser Phe Ser Phe Val Thr Asn Arg Tyr Glu
20 25 30
       Lys Lys Ala Lys Ala Met Gly Asp Val Ala Gly Met Ala Tyr Leu Asp 35 \hspace{1cm} 40 \hspace{1cm} 45
       Ser Net Ser Asn Glu Lys Val Trp Phe Gly Tyr Thr Leu Lys Glu Ala
50 55 60
50
       Gln Ala Gln Gln Ile Gly Leu Gly Leu Asp Leu Lys Gly Gly Het Asn 65 70 75
       Val Ile Leu Lys Leu Asn Ala Ser Asp Leu Leu Arg Asn Leu Ser Asn
85 90 95
55
       Lys Ser Leu Asp Pro Asn Phe Asn Lys Ala Leu Glu Asn Ala Ala Lys 100 105 110
       Ser Thr Glu Gln Ser Asp Phe Ile Asp Ile Phe Val Lys Glu Tyr Arg
115 120 125
       Lys Leu Asp Pro Asn Gly Arg Leu Ala Val Ile Phe Gly Ser Gly Asp
139 135 140
60
      65
       Arg Leu Glu Gly Gln Gly Arg Ile Leu Val Glu Leu Pro Gly Val Lys
195 200 205
       Asp Pro Glu Arg Val Arg Thr Leu Leu Gln Arg Ser Ala Asn Leu Gln 210 220
70
       Phe Trp Arg Thr Tyr Lys Phe Glu Glu Val Ser Gly Asp Leu Ile Ala
225 230 230 235
       Ala Asn Asp Arg Leu Ser Glu Leu Ala Het Asn Asn Thr Asp Ala Thr
245 250 255
75
       Pro Glu Thr Glu Pro Ala Thr Thr Asp Ser Val Ala Ala Thr Ala Asp
```

| | | | | 260 | | | | | 265 | | | | | 270 | | |
|------------|------------|------------|-------------|-----|------------|------------|------------|------------|-----|-------------|------------|------------|------------|-----|------------|------------|
| | Ser | Ala | Ala 275 | Val | Gln | Ala | Val | Ala 280 | Asp | Ser | Ala | Thr | Val 285 | Ala | Gln | ras |
| 5 | Glu | Ala 290 | | Asp | Ala | Thr | Arg 295 | | Asp | Ala | Leu | Phe 300 | | Leu | Leu | Thr |
| | Pro 305 | _ | Asn | Arg | Gly | Gly 310 | | Val | Val | Gly | Val 315 | | Arg | Arg | Ala | Asn 320 |
| | Het | Ala | Gln | Ile | Ser 325 | Glu | Het | Leu | Gln | Gln 330 | Ala | His | Asp | Leu | Lys 335 | Val |
| 10 | | | | 340 | Val | | | | 345 | _ | | - | | 350 | | - |
| | | | 355 | | Lys | | | 360 | | | | | 365 | | | _ |
| 15 | | 370 | | | G1? | | 375 | | | | | 380 | | | | |
| | 385 | | | | lle | 390 | | | | | 395 | | | | | 400 |
| 20 | | | | | Asn 405 | | | | | 410 | | | | | 415 | |
| 20 | | | | 420 | Gly | | | | 425 | | | | | 430 | | |
| | | | 435 | | Asn | | | 440 | | | | | 445 | - | | |
| 25 | | 450 | | | Phe | | 455 | | | | | 460 | | | | |
| | 465 | | | | Lys | 470 | - | | | | 475 | | | | | 480 |
| 20 | | | | | Thr 485 | | | | | 490 | | - | | - | 495 | |
| 30 | | | | 500 | Ala | | | | 505 | | | | | 510 | | |
| | | | 515 | | Pro | | | 520 | | | | | 525 | | | |
| 35 | | 530 | | | Leu | | 535 | | | | | 540 | | | | |
| | 545 | | | | Ala | 550 | | | | | 555 | | | | | 560 |
| 40 | | | | | Ile 565 | | | | | 57 0 | | | | _ | 575 | |
| 40 | | | | 580 | Arg | | | | 585 | | | | | 590 | | |
| | | | 595 | | Ser | | | 600 | | | | | 605 | | | |
| 4 5 | | 610 | | | Thr | | 615 | | | | | 620 | | | | |
| | 625 | | | | Ala | 630 | | | | | 635 | | | | _ | 640 |
| ΕO | | | | | Leu 645 | | | | | 650 | | | | | 655 | |
| 50 | | | | 660 | Thr | | | | 665 | | | | | 670 | | |
| | | | 675 | | Lys | | | 680 | | | | | 685 | | | |
| 55 | | 690 | | | Ala | | 695 | | | | | 700 | | | | |
| | 705 | | _ | _ | Arg | 710 | _ | | | _ | 715 | - | | | | 720 |
| 60 | | | | | Arg 725 | | | | | 730 | | | | | 735 | |
| 60 | | | | 740 | Ile | | | | 745 | | | | _ | 75ů | | |
| | Asn | Tyr | L;∵£ 755 | Ile | Gln | Glu | Glu | Ser 760 | Glu | Glu | Thr | Glu | Ala 765 | Glu | Ile | Thr |
| 65 | | 770 | | | Gln | | 775 | - | - | | - | 780 | | | | |
| | 785 | | | | Leu | 790 | | | | | 795 | | | | | 800 |
| 7 0 | | | | | Asp 805 | | | | | 810 | | | | | 815 | |
| 70 | | | | 820 | Met | | | | 825 | | | | | 830 | | |
| | | | 835 | | Glγ | | | 840 | | | | | 845 | | | _ |
| 75 | Ile | Il∈ 850 | Ala | Leu | Туг | Ala | Leu 855 | Leu | Тгр | Lys | Ile | Leu 860 | Pro | Phe | Thr | Net |

```
Glu lle Asp Gln Asn Phe Ile Ala Ala Ile Leu Ala Ile Ile Gly Tyr
                             870
                                               875
           Leu Asn Asp Thr Val Val Val Phe Asp Arg Ile Arg Glu Thr Met
                        885
                                           890
 5
       Lys Leu Tyr Pro Asn Arg Asp Arg Tyr Gln Val Ile Asn Asp Ala Leu
900 905 910
       Asn Ser Thr Leu Gly Arg Thr Leu Asn Thr Ser Leu Thr Thr Phe Ile
915 920 925
       Val Het Leu Val Ile Phe Ile Phe Gly Gly Ala Thr Met Arg Ser Phe
930 935 940
10
       Thr Phe Ser Ile Leu Leu Gly Ile Val Ile Gly Thr Tyr Ser Thr Leu 945 950 955 960
       Phe Val Ala Thr Pro Leu Ala Tyr Glu Ile Gln Lys Arg Lys Leu Asn
                      965
15
       (2) INFORMATION FOR SEQ ID NO: 424
20
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1017 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
30
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KET: misc_feature
                  (B) LOCATION 1...1017
35
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 424
       Het Lys Arg Met Thr Leu Phe Phe Leu Cys Leu Leu Thr Ser Ile Gly
40
       Trp Ala Met Ala Gln Asn Arg Thr Val Lys Gly Thr Val Ile Ser Ser
20 25 30
       Glu Asp Asn Glu Pro Leu Ile Gly Ala Asn Val Val Val Gly Asn
       Thr Thr Ile Gly Ala Ala Thr Asp Leu Asp Gly Asn Phe Thr Leu Ser 50 60
45
      Val Pro Ala Asn Ala Lys Het Leu Arg Val Ser Tyr Ser Gly Het Thr 65 70 75
      Thr Lys Glu Val Ala Ile Ala Asn Val Met Lys Ile Val Leu Asp Pro
50
      Asp Ser Lys Val Leu Glu Gln Val Val Val Leu Gly Tyr Gly Thr Gly 100 105 110
      Gln Lys Leu Ser Thr Val Ser Gly Ser Val Ala Lys Val Ser Ser Glu
115 120 125
      Lys Leu Ala Glu Lys Pro Val Ala Asn Ile Het Asp Ala Leu Gln Gly
130 135
55
      Gln Val Ala Gly Met Gln Val Het Thr Thr Ser Gly Asp Pro Thr Ala
145 150 155 160
      Val Ala Ser Val Glu Ile His Gly Thr Gly Ser Leu Gly Ala Ser Ser
165 170 175
60
      Ala Pro Leu Tyr Ile Val Asp Gly Het Gln Thr Ser Leu Asp Val Val 180 195 190
      Ala Thr Met Ash Pro Ash Asp Phe Glu Ser Net Ser Val Leu Lys Asp
195 200 205
      Ala Ser Ala Thr Ser Ile Tyr Gly Ala Arg Ala Ala Asn Gly Val Val 210 215 220
65
      Phe lle Gln Thr Lys Lys Gly Lys Het Ser Glu Arg Gly Arg Ile Thr
225 230 235 240
      Phe Asn Ala Ser Tyr Gly Ile Ser Gln Ile Leu Asn Thr Lys Pro Leu
245 250 255
70
      Asp Asn Met Met Thr Gly Asp Glu Leu Leu Asp Phe Gln Val Lys Ala
260 265 270
      Gly Phe Trp Gly Asn Asn Gln Thr Val Gln Lys Val Lys Asp Met Ile
275 280 285
                                                        285
       Leu Ala Gly Ala Glu Asp Leu Tyr Gly Asn Tyr Asp Ser Leu Lys Asp
75
```

| | ~1 | m | | | m) | • | | | | _ | 5 1 | | | | | |
|-----|------------|------------|------------|--------------------------------|------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------------|
| | 305 | Tyr | er?. | LŢS | Thr | 10 Jeu | Phe | Pro | Val | Asp | Phe 315 | Asn | HIS | Asp | Ala | A <i>s</i> p 320 |
| | Trp | Leu | Lys | Ala | Leu | Phe | Lys | Thr | Alā | | | Ser | Gln | Gly | | Ile |
| 5 | Ser | Phe | S∈r | Gly | 325 Gl ₃ | Ser | Gln | Gly | Thr | 330 Ser | T;·r | Tyr | Ala | Ser | 335 Ile | Gly |
| | | | | 340 | | | | | 345 | | | | | 350 | | |
| | ιγι | rne | 355 | GIN | Glu | GT? | riet | 360 | AEG | GIU | PLO | ALa | 365 | rne | гуз | ALG |
| 10 | Tyr | Ser 370 | Glγ | Arg | Leu | Asn | Phe 375 | Glu | Ser | Arg | Ile | Asn 380 | Glu | Trp | Leu | Lys |
| 10 | | | Ala | Asn | Leu | | | Ala | Ile | Ala | Asn | | Arg | Ser | Ala | Asp |
| | 385 Tyr | Phe | Glv | Lvs | туг | 390 Tvr | Met | Glv | Ser | Glv | 395 Thr | Phe | Glv | Val | Leu | 400 Thr |
| 4 = | | | | | 405 | | | - | | 410 | | | | | 415 | |
| 15 | Het | Pro | Arg | 420 | Tyr | Asn | Pro | Phe | 425 | Vāl | Asn | GIŢ | Asp | 430 | Ala | Asp |
| | Val | туг | Tyr 435 | Het | T7.r | Gly | Ala | | Arg | Pro | Ser | Het | Thr 445 | Glu | Pro | T;.r |
| | Ph∈ | Ala | | Het | Arg | Pro | Phe | 440 Ser | Ser | Glu | Ser | His | | Ala | Asn | Val |
| 20 | Asn | 450 Glw | Phe | Ala | Gln | Tle | 455 Thr | Pro | Tle | Line | Gl ·· | 460 Leu | Thr | Leu | Lvs | Ala |
| | 465 | | | | | 470 | | | | _ | 475 | | | | | 480 |
| | Gln | Ala | Gly | Vāl | Asp 485 | Ile | Thr | Asn | Thr | Arg 490 | Thr | Ser | Ser | Lys | Arg 495 | Met |
| 25 | Pro | Asn | Asn | | Tyr | Asp | Ser | Thr | | Leu | Gly | Glu | Arg | | Glu | Arg |
| | Ala | Tyr | Arg | 500 Asp | Val | Ser | Lys | Ser | 505 Phe | Thr | Asn | Thr | Ala | 510 Glu | Tyr | L;:s |
| | Pho | con | 515 | A cr | Glu | 1 | U i a | 520 | 7 | mb a | 7.1 - | Lou | 525 | C1 | шіс | Clu |
| 30 | | 530 | | | | | 535 | - | | | | 540 | | _ | | |
| | Tyr 545 | Ile | Glu | ТУr | Glu | Gly 550 | Asp | Val | Ile | Gl? | Ala 555 | Ser | Ser | Lys | Gly | Phe 560 |
| | | S€r | Asp | $\mathrm{L}\lambda\varepsilon$ | Leu | | Leu | Leu | Ser | | | L's | Thr | Glγ | | |
| 35 | Leu | Ser | Leu | Pro | 565 Glu | His | Arg | Val | Ala | 570 Glu | Tyr | Ala | Tyr | Leu | 575 Ser | Phe |
| | Dhe | Sor | A r-a | 580 Bbc | Asn | m.r. | Clin | Dho | 585 | 1 | T | lie+ | There | 590 | Zan | Dha |
| | | | 595 | | | | | 600 | | | | | 605 | | | |
| 40 | Ser | Val 610 | Arg | Asn | Asp | Gln | Ser 615 | Ser | Arg | Phe | Gly | Ser 620 | Asn | Asn | Arg | Ser |
| | | | Phe | Tyr | Ser | | | Gly | Met | Рþе | | | Tyr | Asn | Lys | |
| | 625 Ile | Gln | Glu | Ser | Asn | 630 Trp | Leu | Ser | Asp | Leu | 635 Arg | Leu | Lys | Met | Ser | 640 Tyr |
| 45 | G1 17 | Thr | Thr | Glv | 645 Asn | Ser | Glu | Tle | Glv | 650 | Tur | Aen | Hic | Gln | 655 | Len |
| | | | | 660 | | | | | 665 | | _ | | | 670 | | |
| | Val | Thr | Val 675 | Asn | Asn | ТУr | Thr | G1u 680 | Asp | Ala | Het | G12 | Leu 685 | Ser | Il∈ | Ser |
| 50 | Thr | Ala 690 | Gly | Asn | Pro | Asp | Leu 695 | Ser | Trp | Glu | Γλε | Gln 700 | Ser | Gln | Phe | Asn |
| 00 | | | Leu | Ala | Ala | Gly | | Phe | Asn | Asn | Arg | | Ser | Ala | Glu | Vāl |
| | 705 Asp | Phe | Tvr | Val | Arg | 710 Thr | Thr | Asn | Aen | Met | 715 Leu | Tle | Asn | Val | Pro | 720 Met |
| 55 | | | | | 725 | | | | | 730 | | | | | 735 | |
| 33 | Fro | ryr | 11e | 3er 740 | Gly | Phe | Phe | Ser | G1n 745 | Tyr | Gln | Asn | Vai | 750 | Ser | Het |
| | Γλ.ε | Asn | Thr 755 | Glγ | Val | Asp | Leu | Ser 760 | Leu | Γλε | Gly | Thr | 11e 765 | туг | Gln | Asn |
| 0.0 | L;:s | Asp | | Aεn | Val | туr | Ala | | Alā | Asn | Phe | Asn | | Asn | Arg | Gln |
| 60 | Glu | 770 Ile | Thr | Lvs | Leu | Phe | 775 Phe | GLV | Len | Asn | Lus | 780 Tur | Het | I.eu | Pro | Asn |
| | 785 | | | | | 790 | | | | | 795 | | | | | 800 |
| | | | | | Trp 805 | | | _ | _ | 810 | | | | | 815 | |
| 65 | Glu | Tyr | Ala | G17 820 | Ile | Asp | r7.2 | Γλε | | G17 | ŗ.'.e | Gln | Leu | | туг | Val |
| | Pro | Glγ | | | Asp | Alā | Asp | | 825 Asn | L,s | Val | Thr | | 830 Ser | Gln | Tyr |
| | | | 835 | | Glu | | | 840 | | | | | 845 | | | |
| 70 | | 850 | | | | | 855 | | - | - | | 860 | | | | |
| | 865 | | | | Ser | 870 | | | | _ | 875 | | | | | 880 |
| | Ala | Asp | Phe | Ala | Tyr 885 | Ile | Val | Gly | Lys | | ilet | Ile | Asn | Asn | | Arg |
| 75 | Tyr | Fhe | Thr | Glu | Asn | Ala | Gl y | G1 ;· | Leu | 890 Het | Gln | Leu | Asn | Γλε | 895 Asp | Lys |
| | | | | | | | | | | | | | | | | |

```
905
       Het Leu Leu Asn Ala Trp Thr Glu Asp Asn Lys Glu Thr Asp Val Pro
915 920 925
       Lys Leu Gly Gln Ser Pro Gln Phe Asp Thr His Leu Leu Glu Asn Ala
930 935 940
       Ser Phe Leu Arg Leu Lys Asn Leu Lys Leu Thr Tyr Val Leu Pro Asn
945 950 960
       Ser Leu Phe Ala Gly Gln Asn Val Ile Gly Gly Ala Arg Val Tyr Leu
965 970 975
10
       Net Ala Arg Asn Leu Leu Thr Val Thr Lys Tyr Lys Gly Phe Asp Pro
980 985 990
       Glu Ala Gly Gly Asn Val Gly Lys Asn Gln Tyr Pro Asn Ser Lys Gln
995 1000 1005
       Tyr Val Ala Gly Ile Gln Leu Ser Phe
15
           1010
       (2) INFORMATION FOR SEQ ID NO: 425
             (i) SEQUENCE CHARACTERISTICS:
20
                   (A) LENGTH: 1014 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
25
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
30
            (ix) FEATURE:
                   (A) MAME/KEY: misc_feature
                   (B) LOCATION 1...1014
35
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:425
       Met Thr Leu Phe Phe Leu Cys Leu Leu Thr Ser Ile Gly Trp Ala Het
       Ala Gln Asn Arg Thr Val Lys Gly Thr Val Ile Ser Ser Glu Asp Asn
20 25 30
40
       Glu Pro Leu Ile Gly Ala Asn Val Val Val Val Gly Asn Thr Thr Ile 35 40
       Gly Ala Ala Thr Asp Leu Asp Gly Asn Phe Thr Leu Ser Val Pro Ala
50 55 60
       Asn Ala Lys Net Leu Arg Val Ser Tyr Ser Gly Het Thr Thr Lys Glu
65 70 75 80
45
       Val Ala Ile Ala As<br/>n Val Met Lys Ile Val Leu Asp Pro Asp Ser L;<br/>s 85 90 95
       Val Leu Glu Gln Val Val Leu Gly Tyr Gly Thr Gly Gln Lys Leu
100 105 110
50
       Ser Thr Val Ser Gly Ser Val Ala Lys Val Ser Ser Glu Lys Leu Ala 115 \phantom{\bigg|} 120 \phantom{\bigg|} 125
       Glu Lys Pro Val Ala Asn Ile Met Asp Ala Leu Gln Gly Gln Val Ala
130 135 140
       Gly Net Gln Val Net Thr Thr Ser Gly Asp Pro Thr Ala Val Ala Ser
145 150 155 160
55
       Val Glu Ile His Gly Thr Gly Ser Leu Gly Ala Ser Ser Ala Pro Leu
165 170 175
       Tyr Ile Val Asp Gly Het Gln Thr Ser Leu Asp Val Val Ala Thr Met
180 185 190
60
       Asn Pro Asn Asp Phe Glu Ser Met Ser Val Leu Lys Asp Ala Ser Ala 195 200 205

Thr Ser Ile Tyr Gly Ala Arg Ala Ala Asn Gly Val Val Phe Ile Gln 210 215 220
65
       Thr Lys Lys Gly Lys Met Ser Glu Arg Gly Arg Ile Thr Phe Asn Ala
225 230 235 240
       Ser Tyr Gly Ile Ser Gln Ile Leu Asn Thr Lys Pro Leu Asp Asn Het
245 250 255
       Het Thr Gly Asp Glu Leu Leu Asp Phe Gln Val Lys Ala Gly Phe Trp 260 265 270
70
       Gly Asn Asn Gln Thr Val Gln Lys Val Lys Asp Het Ile Leu Ala Gly 275 280 285
       Ala Glu Asp Leu Tyr Gly Asn Tyr Asp Ser Leu Lys Asp Glu Tyr Gly 290 295 300
75
       Lys Thr Leu Phe Pro Val Asp Phe Asn His Asp Ala Asp Trp Leu Lys
```

| | 305 | Lou | Dho | Luc | The | 310 | Dro | m1. ~ | C | C1 - | 315 | 7 | T) = | C- × | Dha | 320 |
|------|------------|------------|------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | | | | | 325 | | Pro | | | 330 | | | | | 335 | |
| 5 | | | | 340 | | | Ser | | 345 | | | | | 350 | | |
| | Gln | Glu | G1 y 355 | Het | Ala | Arg | Glu | Pro 360 | Ala | Asn | Phe | Lys | Arg 365 | Туг | Ser | Gly |
| | Arg | Leu 370 | Asn | Phe | Glu | Ser | Arg 375 | Ile | Asn | Glu | Trp | Leu 380 | Γλε | Val | G1γ. | Ala |
| 10 | Asn 385 | Leu | Ser | Gly | Ala | Ile 390 | Ala | Asn | Arg | Arg | Ser 395 | | Asp | Tyr | Phe | Gly 400 |
| | | Tyr | Tyr | Het | Gly 405 | | Gly | Thr | Phe | Gly 410 | | Leu | Thr | Met | Pro 415 | |
| 15 | Туr | Туr | Asn | Pro 420 | | Asp | Val | Asn | Gly 425 | | Leu | Ala | Asp | Val 430 | | туг |
| 10 | Het | Tyr | Gl _y 435 | | Thr | Arg | Pro | | | Thr | Glu | Pro | | | Ala | Γ7.ε |
| | Het | | | Phe | Ser | Ser | Glu | 440 Ser | His | Gln | Ala | | 445 Val | Asn | G1y | Phe |
| 20 | | 450 Gln | Ile | Thr | Pro | | 455 Lys | Glγ | Leu | Thr | | 460 Lys | Alā | Gln | Ala | |
| | 465 Val | Asp | lle | Thr | | 470 Thr | Arg | Thr | Ser | | 475 Lys | Arg | Net | Pro | | 480 Asn |
| or | Pro | туг | Asp | | 485 Thr | Pro | Leu | Glγ | | 490 Arg | Arg | Glu | Arg | Ala | 495 Tyr | Arg |
| 25 | Asp | Val | Ser | 500 Lys | Ser | Phe | Thr | Asn | 505 Thr | Ala | Glu | Tyr | Lys | 510 Phe | Ser | Ile |
| | Asp | Glu | 515 Lys | Hiε | Asp | Leu | Thr | 520 Ala | Leu | het | Gly | His | 525 Glu | Tyr | Ile | Glu |
| 30 | Tyr | 530 Glu | Gly | Asp | Val | Ile | 535 Gly | Ala | Ser | Ser | Lys | 540 Gly | Phe | Glu | Ser | Asp |
| | 545 | | | | | 550 | Gln | | | | 555 | | | | | 560 |
| | | | | | 565 | | Glu | | | 570 | | | | | 575 | |
| 35 | | | | 580 | | | Lys | | 585 | | | | | 590 | | |
| | | | 595 | | | | Phe | 600 | | | | | 605 | | | |
| 40 | | 610 | | | | | 615 | | | | | 620 | | | | |
| 40 | 625 | | | | | 630 | Phe | | | | 635 | | | | | 640 |
| | | | | | 645 | | Leu - | | | 650 | | | | | 655 | |
| 45 | | | | 660 | | | Asn | - | 665 | | | | | 670 | | |
| | Asn | Asn | Tyr 675 | Thr | Glu | Asp | Ala | Het 680 | Gly | Leu | Ser | Ile | Ser 685 | Thr | Ala | Gly |
| | Aεn | Pro 690 | Asp | Leu | Ser | Trp | Glu 695 | r?.e | Gln | Ser | Gln | Phe 700 | Asn | Phe | G1; | Leu |
| 50 | Ala 705 | Ala | GI 7. | Ala | Phe | Asn 710 | Asn | Arg | Leu | Ser | Ala 715 | Glu | Val | Asp | Phe | Туг 720 |
| | | | | | 725 | | Met | | | 730 | | | | | 735 | |
| 55 | | | | 740 | | | Туг | | 745 | | | | | 750 | | |
| | G1 y | Val | Asp 755 | Leu | Ser | Leu | Lys | Gly 760 | Thr | Ile | Туr | Gln | Asn 765 | Lys | Asp | Trp |
| | Asn | Val 770 | Tyr | Ala | Ser | Ala | Asn 775 | | Asn | Туr | Asn | Arg 780 | | Glu | Ile | Thr |
| 60 | Lys 785 | Leu | Phe | Phe | Gly | Leu 790 | Asn | r?.e | Tyr | l·let | Leu 795 | | Asn | Thr | Gl? | Thr 800 |
| | | Trp | Glu | Ile | Gly 805 | | Pro | Asn | Ser | Phe 810 | | Het | Ala | Glu | Tyr 815 | |
| 65 | Gly | Ile | Asp | Lys 820 | | Thr | G17. | Γλ.ε | Gln 825 | | Trp | Туг | Val | | | Gln |
| 00 | Val | Λsp | | | Gly | Asn | Lyε | | | Thr | Ser | Gln | | 830 Ser | Ala | Asp |
| | Leu | | 835 Thr | Arg | Ile | Asp | L;:s | 840 Ser | Val | Thr | Pro | | 845 Ile | Thr | Gly | Gly |
| 70 | | 850 Ser | Leu | Gly | Ala | | 855 Trp | Lys | Glу | Leu | | 860 Leu | Asp | Ala | Asp | |
| | 865 Ala | Туг | He | Val | | 870 Lys | Trp | Het | Ile | | 875 Asn | Asp | Arg | туг | | 880 Thr |
| r) F | Glu | Asn | Ala | Gly | 885 Gly | Leu | Het | Gln | Leu | 890 Asn | L;/s | Asp | Lys | Met | 895 Leu | Leu |
| 75 | | | | 900 | | | | | 905 | | | - | | 910 | | |

```
Asn Ala Trp Thr Glu Asp Asn Lys Glu Thr Asp Val Pro Lys Leu Gly 915 920 925
       Gln Ser Pro Gln Phe Asp Thr His Leu Leu Glu Asn Ala Ser Phe Leu
930 935 940
       Arg Leu Lys Asn Leu Lys Leu Thr Tyr Val Leu Pro Asn Ser Leu Phe 945 950 955 960
 5
       Ala Gly Gln Asn Val Ile Gly Gly Ala Arg Val Tyr Leu Het Ala Arg
965 970 975
       Asn Leu Leu Thr Val Thr Lys Tyr Lys Gly Phe Asp Pro Glu Ala Gly 980 985 990
Gly Asn Val Gly Lys Asn Gln Tyr Pro Asn Ser Lys Gln Tyr Val Ala
10
                995
       Gly Ile Gln Leu Ser Phe
15
       (2) INFORMATION FOR SEQ ID NO: 426
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 821 amino acids
20
                   (P) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
25
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
30
            (ix) FEATURE:
                   (A) HAME/KEY: misc_feature
                   (B) LOCATION 1...821
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 426
35
       Het Lys Lys Lys Asn Phe Leu Leu Eu Gly Ile Phe Val Ala Leu Leu 1 5 10 15
       Thr Phe Ile Gly Ser Met Gln Ala Gln Gln Ala Lys Asp Tyr Phe Asn 20 25 30
40
       Phe Asp Glu Arg Gly Glu Ala Tyr Phe Ser Phe Lys Val Pro Asp Arg 35 40 45
       Ala Val Leu Gln Glu Leu Ala Leu Ile Met Ser Ile Asp Glu Phe Asp
50 55 60
       Pro Val Thr Asn Glu Ala Ile Ala Tyr Ala Ser Glu Glu Glu Phe Glu
65 70 80
45
       Ala Phe Leu Arg Tyr Gly Leu Lys Pro Thr Phe Leu Thr Pro Pro Ser
85 90
       Het Gln Arg Ala Val Glu Het Phe Asp Tyr Arg Ser Gly Glu Lys Tyr
100 105
       Glu Trp Asn Ala Tyr Pro Thr Tyr Glu Ala Tyr Ile Ser Het Met Glu
115 120 125
50
       Glu Phe Gln Thr Lys Tyr Pro Ser Leu Cys Thr Thr Ser Val Ile Gly
130 135 140
       Lys Ser Val Lys Asp Arg Lys Leu Met Ile Cys Lys Leu Thr Ser Ser
145 150 155 160
55
       Ala Asn Thr Gly Lys Lys Pro Arg Val Leu Tyr Thr Ser Thr Net His
165 170 175
       Gly Asp Glu Thr Thr Gly Tyr Val Val Leu Leu Arg Leu Ile Asp His
180 185 190
60
       Leu Leu Ser Asn Tyr Glu Ser Asp Pro Arg Ile Lys Asn Ile Leu Asp
195 200 205
       Lys Thr Glu Val Trp Ile Cys Pro Leu Thr Asn Pro Asp Gly Ala Tyr 210 220
       Arg Ala Gly Asn His Thr Val Gln Gly Ala Thr Arg Tyr Asn Ala Asn
225 230 235 240
65
       Asn Val Asp Leu Asn Arg Asn The Lys Asp Asp Val Ala Gly Asp His
245 255
       Pro Asp Gl; Lys Pro Trp Gln Pro Glu Ala Thr Ala Phe Met Asp Leu
260 265 270
70
       Glu Gly Asn Thr Ser Phe Val Leu Gly Ala Asn Ile His Gly Gly Thr
275 280 285
       Glu Val Val Asn Tyr Pro Trp Asp Asn Lys Lys Glu Arg His Ala Asp
290 295 300
                                                        300
       Asp Glu Trp Tyr Lys Leu Ile Ser Arg Asn Tyr Ala Ala Ala Cys Gln
305 310 315
75
```

```
Ser Ile Ser Ala Ser Tyr Met Thr Ser Glu Thr Asn Ser Gly Ile Ile 325 \hspace{1cm} 330 \hspace{1cm} 335
       Asn Gly Ser Asp Trp Tyr Val Ile Arg Gly Ser Arg Gln Asp Asn Ala
340 345 350
       Asn Tyr Phe His Arg Leu Arg Glu Ile Thr Leu Glu Ile Ser Asn Thr 355 360 365
       Lys Leu Val Pro Ala Ser Gln Leu Pro Lys Tyr Trp Asn Leu Asn Lys
370 375 380
       Glu Ser Leu Leu Ala Leu Ile Glu Glu Ser Leu Tyr Gly Ile His Gly
395 390 400
10
       Thr Val Thr Ser Ala Ala Asn Gly Gln Pro Leu Lys Cys Gln Ile Leu
405 410 415
       Ile Glu Asn His Asp Lys Arg Asn Ser Asp Val Tyr Ser Asp Ala Thr 420 425 430
       Thr Gly Tyr Tyr Val Arg Pro Ile Lys Ala Gly Thr Tyr Thr Val Lys 435 440 445
15
       Tyr Lys Ala Glu Gly Tyr Pro Glu Ala Thr Arg Thr Ile Thr Ile Lys
450 455 460
       Asp Lys Glu Thr Val Ile Net Asp Ile Ala Leu Gly Asn Ser Val Pro
465 470 470 480
20
       Leu Pro Val Pro Asp Phe Thr Ala Ser Pro Net Thr Ile Ser Val Gly 485 490 495
       Glu Ser Val Gln Phe Gln Asp Gln Thr Thr Asn Asn Pro Thr Asn Trp
500 505 510
25
       Glu Trp Thr Phe Glu Gly Gly Gln Pro Ala Net Ser Thr Glu Gin Asn
515 520 525
       Pro Leu Val Ser Tyr Ser His Pro Gly Gln Tyr Asp Val Thr Leu Lys
530 540
       Val Trp Asn Ala Ser Gly Ser Asn Thr Ile Thr Lys Glu Lys Phe Ile
545 550 555 560
30
       Thr Val Asn Ala Val Met Pro Val Ala Glu Phe Val Gly Thr Pro Thr 565 570 575
                       565
       Glu Ile Glu Glu Gly Gln Thr Val Ser Phe Gln Asn Gln Ser Thr Asn
580 585 590
35
       Ala Thr Asn Tyr Val Trp Ile Phe Asp Gly Gly Thr Pro Ala Thr Ser
595 600 605
             595
       Glu Asp Glu Asn Pro Thr Val Leu Tyr Ser Lys Ala Gly Gln Tyr Asp
610 620
       Val Thr Leu Lys Ala Ile Ser Ala Ser Gly Glu Thr Val Lys Thr Lys
625 630 640
40
       Glu Lys Tyr Ile Thr Val Lys Lys Ala Pro Val Pro Ala Pro Val Ala
645 650 655
       Asp Phe Glu Gly Thr Pro Arg Lys Val Lys Lys Gly Glu Thr Val Thr
660 665 670
       Phe Lys Asp Leu Ser Thr Asn Asn Pro Thr Ser Trp Leu Trp Val Phe 675 680 685
45
       Glu Gly Gly Ser Fro Ala Thr Ser Thr Glu Gln Asn Pro Val Val Thr 690 700
       Tyr Asn Glu Thr Gly Lys Tyr Asp Val Gln Leu Thr Ala Thr Asn Glu 705 710 715 720
50
       Gly Gly Ser Asn Val Lys Lys Ala Glu Asp Tyr Ile Glu Val Ile Leu
725 730 735
       Asp Asp Ser Val Glu Asp Ile Val Ala Gln Thr Gly Ile Val Ile Arg
740 745 750
55
       Pro Gln Asn Gly Thr Lys Gln Ile Leu Ile Glu Ala Asn Ala Ala Ile
755 760 765
       Lys Ala Ile Val Leu Tyr Asp Ile Asn Gly Arg Val Val Leu Lys Thr
770 780
       Thr Pro Asn Gln Leu Arg Ser Thr Val Asp Leu Ser Ile Leu Pro Glu
785 790 795 800
60
       Gly Ile Tyr Thr Ile Asn Ile Lys Thr Glu Lys Ser Ala Arg Thr Glu
65
       (2) INFORMATION FOR SEQ ID NO: 427
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 76 amino acids
70
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
75
           (iii) HYPOTHETICAL: YES
```

```
(vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
 5
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...76
           (mi) SEQUENCE DESCRIPTION: SEQ ID NO: 427
10
      Het Arg Leu lle Lys Ala Phe Leu Val Gln Leu Leu Leu Pro Ile
      Phe Phe Tyr Lys Arg Phe Ile Ser Pro Leu Thr Pro Pro Ser Cys Arg 20 25 30
      Phe Thr Pro Ser Cys Ser Ser Tyr Ala Ile Glu Ala Leu Arg Lys Tyr 35 40 45
15
      Gly Pro Gly Lys Gly Leu Leu Ser Ile Lys Arg Ile Leu Arg Cys
50 55 60
      His Pro Trp Gly Gly Ser Gly Tyr Asp Pro Val Pro
20
      (2) INFORMATION FOR SEQ ID NO:428
            (i) SEQUENCE CHARACTERISTICS:
25
                  (A) LENGTH: 859 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
30
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
35
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...859
40
           (mi) SEQUENCE DESCRIPTION: SEQ ID NO:428
      Het Ala Tyr Asp the Thr Gln Thr Phe Arg Asn Ser Leu Glu Tyr Ser
      Tyr Gln Glu Ala Thr Arg Leu Gly Val Val Ala Val Thr Gln Asp Met 20 25 30
45
      Leu Val Leu Gly Ile Ile Arg Asp Gly Asp Asn Gly Ala Ile Asp Ile 35 40 45
      Met Arg His Tyr Gly Ile Asn Leu Tyr Glu Leu Lys Arg Leu Ile Glu
50 55 60
50
       Leu Glu Ala Ile Ala Glu Ser Leu Pro Ala Ser Pro Glu Gly Ser Pro
65 70 75 80
      The Phe Thr Pro Ser Ala Arg Glu Ala Ile Asp Asp Ala Thr Asp Ile
85 90 95
       Cys Ala Asp Met Glu Asp Glu Ala Val Ser Pro Val His Leu Leu
100 105 110
55
       Ser Ile Leu Asn Ser Thr Gln Glu Ser Leu Val Gln Lys Ile Phe Met
115 120 125
       Lys Gln Gly Ile Lys Tyr Asp Thr Ile Leu Ser Asp Tyr Phe Gly Gln 130 $135$
      Arg Asn Pro Ser Glu Gly Lys Ser Pro Ser Glu Met Glu Ile Leu Asp
145 150 160
60
       Gly Tyr Gln Asp Asp Asp Phe Asp Asp Glu Glu Asp Glu Ser Ser Pro
165 170 175
       Pro Ser Gly Asn Ser Gly Thr Gly Gly Gly Ser Gly Asp Ala Pro Glu
180 185 190
65
       Gln Asn Thr Gly Gly Gly Asp Thr Thr Thr Thr Arg Ser Gly Gly
195 200 205
       Asp Thr Pro Ala Leu Asp Thr Phe Gly Thr Asp Ile Thr Ala Net Ala 210 215 220
70
       Ala Ala Gly Lys Leu Asp Pro Val Val Gly Arg Glu Gln Glu Ile Glu
225 230 236 240
       Arg Val Ile Gln Ile Leu Ser Arg Arg Lys Lys Asn Asn Pro Val Leu 245 250 255
       Ile Gly Glu Pro Gly Val Gly Lys Ser Ala Ile Val Glu Gly Leu Ala
260 265 270
75
```

| | Glu | Arg | | Val | Asn | Arg | Lys | | Ser | Arg | Ile | Pen | | Asp | Lys | Arg |
|-----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------|------------|------------|------------|------------|-------------|
| | Ile | He | 275 Ser | Leu | Asp | Leu | Ala | 280 Gln | llet | Val | Ala | Gly | 285 Thr | Lγε | Туr | Arg |
| 5 | Gly | 290 Gln | Phe | Glu | Glu | Arg | 295 Leu | Lys | Ala | Val | Leu | 300 Asp | Glu | Leu | Lys | L <i>ys</i> |
| | 305 Asn | Pro | Gln | Ile | Ile | 310 Leu | Phe | Ile | Asp | Glu | 315 Ile | His | Thr | Ile | Val | 320 Gly |
| 40 | Ala | Gly | Ser | Ala | 325 Ala | Gly | Ser | Het | | 330 Thr | Alā | Asn | Met | | 335 Lys | Pro |
| 10 | Ala | Leu | | 340 Arg | Gly | Gln | Vāl | | 345 Cys | Ile | G1; | Ala | | 350 Thr | Leu | Asp |
| | Glu | | 355 Arg | Lys | Asn | Ile | | 360 Lys | Asp | Gly | Ala | | 365 Glu | Arg | Arg | Phe |
| 15 | Gln 385 | 370 Lys | Val | Pro | Ile | Ala 390 | 375 Pro | Ser | Thr | Ala | | 380 Glu | Thr | Leu | Thr | |
| | | Gln | Asn | Ile | Lys 405 | | Lys | туг | Glu | Asp 410 | 395 Tyr | His | Glγ | Val | Arg 415 | 400 Туг |
| 20 | Thr | Asp | Glu | Ala 420 | | Lys | Ala | Ala | Val 425 | | Leu | Thr | Asp | Arg 430 | | Val |
| | Ser | Asp | Arg 435 | Phe | Phe | Pro | Asp | Lys 440 | | Ile | Asp | Ala | Het 445 | | Glu | Ala |
| | Gly | Ala 450 | | Val | His | Ile | Thr 455 | | Val | Val | Alā | Pro 460 | | Glu | Ile | Glu |
| 25 | Ile 465 | Leu | Glu | Ala | Glu | Leu 470 | | Ser | Val | Arg | Glu 475 | | Lys | Leu | Ser | Ala 480 |
| | Val | Γλ.ε | Ala | Gln | Asn 485 | Tyr | Glu | Leu | Ala | Ala 490 | Ser | Phe | Arg | Asp | Gln 495 | |
| 30 | Arg | Arg | Thr | Gln 500 | Gln | Gln | Ile | Ala | Glu 505 | Glu | Γλε | Lys | Lys | Trp 510 | Glu | Glu |
| | Gln | Met | Ser 515 | Lys | His | Arg | Glu | Thr 520 | Vāl | Asp | Glu | Asn | Val 525 | Val | Ala | His |
| 25 | | 530 | | Leu | | | 535 | | | | | 540 | | | | _ |
| 35 | 545 | | | Arg | | 550 | | | | | 555 | | | | | 560 |
| | | | | Asp | 565 | | | | | 570 | | | | | 575 | |
| 40 | | | | Gly 580 | | | | | 585 | | | | | 590 | | |
| | | | 595 | Pro | | | | 600 | | | | | 605 | | | |
| 45 | | 610 | | Leu | | | 615 | | | | | 620 | | | | |
| 40 | 625 | | | Met | | 630 | | | | | 635 | | | _ | | 640 |
| | | | | Val Pro | 645 | | | | _ | 650 | | | | | 655 | |
| 50 | | | | 660 Val | | | | | 665 | | | | | 670 | - | |
| | | | 675 | Ser | | | | 680 | | | | | 685 | | | |
| 55 | | 690 | | Ser | | | 695 | | | | | 700 | | | | |
| | 705 | | | Phe | | 710 | | | | | 715 | | | | | 720 |
| | | | | Val | 725 | | | | | 730 | | | | | 735 | |
| 60 | | | | 740 Arg | | | | | 745 | | | | | 750 | | |
| | Thr | Glu | 755 Ile | Arg | Arg | Met | Val | 760 Asp | Ile | Glu | Leu | L)'s | 765 Ala | Vál | Leu | Ala |
| 65 | | 770 Ile | His | Arg | Ala | | 775 Tyr | Asp | Leu | Val | Leu | 780 Thr | Asp | Glu | Ala | Lys |
| | 785 Asp | Vāl | Ile | Ala | | 790 Lys | Glγ | Tyr | Asp | | 795 Gln | T).r | GΙλ | Ala | | B00 Fro |
| 70 | Leu | L;;s | Arg | Thr | 805 Leu | G1n | Asn | Glu | | 810 Glu | Asp | Arg | Leu | | 815 Asp | Leu |
| , 0 | Ile | Leu | | 820 Gly | Gln | Ile | Glu | | 825 Gly | Gln | T'h <i>r</i> | Leu | | Leu Leu | Ser | Ala |
| | Arg | Asp 850 | 835 Gl; | Glu | Ile | Ile | | 840 Gln | Glu | Gln | Ala | | 845 | | | |
| 75 | | 0.50 | | | | | 855 | | | | | | | | | |

```
(2) INFORMATION FOR SEQ ID NO: 429
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 417 amino acids
 5
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
10
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
15
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...417
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:429
20
       Ser Gly Ala Glu Ile Ala Phe Leu Ser Ser Asp Lys Leu Arg Leu Glu 20 25 30
25
       Leu Asp Arg Ash Arg Gly Asp Leu Thr Gly Arg Ala Leu Ash Leu Leu 35 40 45
       Tyr Arg His Pro Asp Gln Leu Val Thr Thr Leu Leu Val Gly Asn Asn 50 60
       lle Val Leu Val Val Tyr Gly Leu Leu Het Ala Gly Leu Leu Ala Ala 65 \phantom{000}70\phantom{000} 75 \phantom{0000}80\phantom{000}
30
       Pro Leu Ala Gln Trp Ile Asp Asn Asp Ala Net Ile Val Val Leu Gln
85 90 95
       Ser Val Leu Ser Thr Ile Ile Ile Leu Phe Thr Gly Glu Phe Leu Pro
100 105 110
35
       Lys Ala Ile Phe Lys Thr Asn Ala Asn Net Met Met Arg Val Phe Ala 115 120 125
       Leu Fro Ile Val Ala Ile Tyr Tyr Leu Leu Tyr Pro Leu Ser Lys Leu
130 135 140
       Phe Thr Gly Leu Ser Arg Ser Phe Ile Arg Leu Val Asp Lys Asn Tyr
145 150 160
40
       Val Pro Thr Thr Val Gly Leu Gly Arg Val Asp Leu Asp His Tyr Leu
165 170 175
       Ala Glu Asn Het Ser Gly Glu Asn Glu Gln Asn Asp Leu Thr Thr Glu
180 185 190
45
       Val Lys Ile Ile Gin Asn Ala Leu Asp Phe Ser Gly Ile Gin Val Arg
195 200 205
       Asp Cys Het Ile Pro Arg Asn Glu Het Ile Ala Cys Glu Leu Gln Thr 210 215 220
       Asp Ile Glu Val Leu Lys Thr Thr Phe Ile Asp Thr Gly Leu Ser Lys 225 230 235 240
50
       The The The Tyr Arg Gln Asn The Asp Asp Val Val Gly Tyr The His 245 250 255
       Ser Ser Glu Met Phe Arg Gly Gln Asp Trp Gln Lys Arg Ile Asn Thr
260 265 270
55
       Thr Val Phe Val Pro Glu Ser Het Tyr Ala Asn Lys Leu Het Arg Leu
275 280 285
       Leu Het Gln Arg Lys Lys Ser Ile Ala Ile Val Ile Asp Glu Leu Gly 290 295 300
       Gly Thr Ala Gly Net Val Thr Leu Glu Asp Leu Val Glu Glu Ile Phe 305 310 315
60
       Gly Asp Ile Glu Asp Glu His Asp Thr Arg Lys Ile Ile Ala Lys Gln
325 330 335
       Leu Gly Pro His Thr Tyr Leu Val Ser Gly Arg Het Glu Ile Asp Asp 340 345
65
       Val Asn Glu Arg Phe Gly Leu Ser Leu Pro Glu Ser Asp Asp Tyr Leu
355 360 365
       Thr Val Ala Gly Phe Ile Leu Asn Ser His Gln Asn Ile Pro Gln Ala
370 375 380
       Asn Glu Val Val Glu Ile Ala Pro Tyr Thr Phe Thr Ile Leu Arg Ser
385 390 395 400
70
       Ser Ser Thr Lys Ile Glu Leu Val Lys Net Ser Ile Asp Asp Gln Ser
```

75

```
(2) INFORMATION FOR SEQ ID NO:430
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 293 amino acids
 5
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
10
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
15
           (ix) FEATURE:
                  (A) NAME/KET: misc_feature
                  (B) LOCATION 1...293
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:430
20
      Net Lys Gln Asn Tyr Phe Lys Arg Val Cys Ser Leu Leu Trp Leu Val 1 5 10 15
       Leu Pro Met Leu Ile Met Pro Leu Glu Val Ala Ala Gln Glu Ile Ile
                                          25
25
       Pro Asn Glu Glu Val Leu Glu Ser Leu Thr Phe Val Ala Pro Val Glu
                                     4.0
       Glu Thr Asp Ala Ile Glu Ala Glu Val Glu Ala Leu Gln Glu Ile Val
50 55 60
       Ala Thr Glu Glu Ile Ala Glu Gln Ala Val Arg Ser Tyr Thr Tyr Thr
65 70 80
30
       Val T;r Arg Asp Gly Val Lys Ile Ala Ser Gl; Leu Thr Glu Pro Thr 85 90 95
       Phe Leu Asp Glu Asp Val Pro Ala Gl; Glu His Thr Tyr Cys Val Glu
100 105 110
      Val Gln Gly Gly Val Ser Asp Lys Val Cys Val Asp Val Glu
115 120 125

Val Lys Asp Phe Lys Pro Val Thr Asn Leu Thr Gly Thr Ala Ser Asn
130 135 140
35
       Asp Glu Val Ser Leu Asp Trp Asp Gly Val Glu Glu Lys Ala Glu Glu
145 150 155 160
40
       Pro Ala Ser Asp Lys Ala Val Ser Tyr Asn Val Tyr Lys Asn Gly Thr
165 170 175
       Leu Ile Gly Asn Thr Ala Glu Thr His Tyr Val Glu Thr Gly Val Ala
180 185 190
45
       Asn Gly Thr Tyr Ile Tyr Glu Val Glu Val Lys Tyr Pro Asp Gly Val
195 200 205
               195
       Ser Pro Lys Val Ala Val Thr Val Thr Val Thr Asn Ser Ser Leu Ser
                              215
       Asn Val Asp Gly Gln Ala Pro Tyr Thr Leu Arg Val Glu Gly Lys Lys
225 230 235 240
50
       Ile Ile Ala Glu Ala His Gly Met Ile Thr Leu Tyr Asp Ile Asn Gly
                       245
                                             250
      Arg Thr Val Ala Val Ala Pro Asn Arg Leu Glu Tyr Met Ala Gln Thr 260 265 270
55
       Gly Phe Tyr Ala Val Arg Phe Asp Val Gly Asn Lys His His Val Ser
       Lys Ile Gln Val Arg
60
       (2) INFORMATION FOR SEQ ID NO:431
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 312 amino acids
                  (B) TYPE: amino acid
65
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
70
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
75
                 (A) NAME/KEY: misc_feature
```

(B) LOCATION 1...312

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:431
      Met Ile Pro Leu Ser Glu Ser Phe Glu Ser Gly Ile Pro Ala Ile Trp
                                           10
      25
      Asn Phe Thr Gly Gln Ser Gly Leu Cys Val Ser Ser Ala Ser Tyr Ile
10
                                  40
      Gly Gly Val Gly Ala Leu Thr Pro Asp Asn Tyr Leu Ile Thr Pro Glu
                            55
      Leu Lys Leu Pro Thr Asp Ala Leu Val Glu Ile Ile Tyr Trp Val Cys
                                             75
      Thr Gln Asp Leu Thr Ala Pro Ser Glu His Tyr Ala Val Tyr Ser Ser
85 90
15
      Ser Thr Gly Asn Asn Ala Ala Asp Phe Val Asn Leu Leu Tyr Glu Glu
100 105 110
      Thr Leu Thr Ala Lys Arg Ile Gln Ser Pro Glu Leu Ile Arg Gly Asn
115 120 125
20
      Arg Thr Gin Gly Val Trp Tyr Gin Arg Lys Val Val Leu Pro Asn Asp
130 135
      Thr Lys Tyr Val Ala Phe Arg His Phe Asn Ser Thr Asp Asn Phe Trp
145 150 160
      Leu Asn Leu Asp Glu Val Ser Ile Leu Tyr Thr Pro Leu Pro Arg Arg
165 170 175
25
      Ala Pro Cys Pro His Pro Gly Gly Tyr Thr Tyr Ser Val Phe Arg Asp
180 185 190
      Gly Gln Lys Ile Ala Ser Gly Leu Ser Ala Leu Ala Tyr Ile Asp Thr
195 200 205
30
      Asp Val Pro Tyr Gly Thr Gln Asp Tyr Cys Val Gln Val Asn Tyr Leu 210 215 220 220 220 215 220 235 240
      Ala Asn Ile Tyr Gly Ala Asp Lys Pro Phe Ala Leu Thr Val Val Gly 245 250 255
35
      Lys Thr Ile Val Ala Ser Ala Phe Lys Gly Glu Ile Thr Leu Tyr Asp
260 265 270
      lle Arg Gly Arg Leu Ile Ala Ser Gly Cys Asp Thr Leu Arg Tyr Lys
275 280 285
40
      Ala Glu Asn Gly Phe Tyr Leu Ile Lys Ile Gln Val Asn Gly Thr Val
                              295
         290
      Tyr Thr Glu Lys Ile Gln Ile Gln
45
      (2) INFORMATION FOR SEQ ID NO:432
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 843 amino acids
50
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: Porphyromonas gingivalis
60
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...843
           (::i) SEQUENCE DESCRIPTION: SEQ ID NO:432
65
      Net Lys Lys Ser Phe Leu Leu Ala Ile Val Net Leu Phe Gly Ile Ala
                                            10
       Het Glm Gly His Ser Ala Pro Val Thr Lys Glu Arg Ala Leu Ser Leu
      Ala Arg Leu Ala Leu Arg Gln Val Ser Leu Arg Met Gly Gln Thr Ala
35 40 45
70
       Val Ser Asp Lys Ile Ser Ile Asp Tyr Val Tyr Arg Gln Gly Asp Ala
                             55
                                                   60
       Glu Arg Gly Ile Thr Ser Gln Glu Glu Gly Ser Pro Ala Tyr Phe Tyr
75
```

| | Val | Ala | A≲n | Arg | Gly 85 | Asn | A≋n | Glu | Gly | Туг 90 | Ala | Leu | Val | Ala | Ala 95 | Asp |
|------------|------------|------------|------------|--------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Asp | Arg | Ile | Pro 100 | Thr | lle | Leu | Ala | Tyr 105 | Ser | Pro | Ile | G13 | Arg 110 | Phe | Asp |
| 5 | llet | Asp | Ser 115 | Net | Pro | Asp | Asn | Leu 120 | Arg | Met | Trp | Leu | Gln 125 | Ile | Туr | Asp |
| | Gln | Glu 130 | Ile | Gly | Leu | Ile | Leu 135 | Ser | Gly | Γλ.ε | Ala | Gln 140 | Leu | Asn | Glu | Glu |
| 10 | 11e 145 | Leu | Arg | Thr | Glu | Gly 150 | Val | Pro | Ala | Glu | Val 155 | Hiε | Ala | Leu | Net | Asp 160 |
| | Asn | Glγ | His | Phe | Ala 165 | Asn | Asp | Pro | Met | Arg 170 | Trp | Asn | Gln | G1 ?. | Tyr 175 | Pro |
| | Trp | Asn | Asn | L;s 180 | Glu | Pro | Leu | Leu | Pro 185 | Asn | G1; | Asn | His | Ala 190 | Tyr | Thr |
| 15 | Glγ | Cys | Val 195 | Ala | Thr | Ala | Ala | Ala 200 | Gln | Ile | Het | Arg | Tyr 205 | His | Ser | Trp |
| | Pro | Leu 210 | Gln | Gly | Glu | Gly | Ser 215 | Phe | Asp | T; r | His | Ala 220 | Gly | Ser | Leu | Val |
| 20 | Gly 225 | Asrı | Trp | Ser | Gly | Thr 230 | Phe | Gly | Glu | Het | Tyr 235 | A≥p | Trp | Ile | Asn | Met 240 |
| | Pro | Gly | Asn | Pro | Asp 245 | Leu | Asp | Asn | Leu | Thr 250 | Gln | Ser | Gln | Val | Asp 255 | Ala |
| | T7.r | Ala | Thr | Leu 260 | Иet | Arg | Asp | Val | Ser 265 | Ala | Ser | Vāl | Ser | Het 270 | Ser | Phe |
| 25 | Tyr | Glu | Asn 275 | Gly | Ser | Gly | Thr | Tyr 280 | Ser | Val | T;r | Val | Val 285 | Glγ | Ala | Leu |
| | Arg | Asn 290 | Asn | Phe | Arg | Tyr | Lys 295 | Arg | Ser | Leu | Gln | Leu 300 | His | Val | Arg | Ala |
| 30 | 305 | | | Ser | | 310 | | | | | 315 | | | | | 320 |
| | Ser | G1 y | Arg | Pro | Val 325 | Tyr | Туг | Ala | Gly | A⊊n 330 | Asn | Gln | Ser | Ile | Gly 335 | His |
| 0.5 | | | | C <i>∵∈</i> 340 | | | | | 345 | | | | | 350 | | |
| 35 | Trp | Gly | Trp 355 | Gly | Gly | Val | Ser | Asn 360 | Gly | Phe | Tyr | Lys | Leu 365 | Thr | Leu | Leu |
| | | 370 | | Ser | | | 375 | | | | | 380 | | | | |
| 40 | 385 | | | Ile | | 390 | | | | | 395 | | | | | 400 |
| | | | | Asp | 405 | | | | | 410 | | | | | 415 | |
| 45 | | | | Ser 420 | | | | | 425 | | | | | 430 | | |
| 45 | | | 435 | Glu | | | | 440 | | | | | 445 | | | |
| | | 450 | | Glu | | | 455 | | | | | 460 | | | | |
| 50 | 465 | | | Tyr | | 470 | | | | | 475 | | | | | 480 |
| | | | | Gln | 485 | | | | | 490 | | | _ | | 495 | |
| 55 | | | | Gln 500 | | | | | 505 | | | | | 510 | | |
| J J | | | 515 | Lys | | | | 520 | | | | | 525 | | | |
| | | 530 | | Asn | | _ | 535 | | | | | 540 | | | | |
| 60 | 545 | | | Asn | | 550 | | | | | 555 | | | | | 560 |
| | | | | Pro | 565 | | | | | 570 | | | | | 575 | |
| 65 | | | | Thr 580 | | | | | 585 | | | | | 590 | | |
| 00 | | | 595 | Gly | | | | 600 | | | | | 605 | | | |
| | | 610 | | Leu | | | 615 | | | | | 620 | | | | |
| 70 | 625 | | | Gln | | 630 | | | | | 635 | | | | | 640 |
| | | | | Thr | 645 | | | | | 650 | | | | _ | 655 | |
| 75 | | | | Thr 660 | | | | | 665 | | | | | 670 | | |
| , 5 | vab | rne | ınr | Leu | r?/s | ASD | геп | ÕΤ. | ren | Pro | Phe | Asn | GTA | Glu | ren | val |

```
680
                 675

      Val Val Phe Arg Gln Thr Gln Ser Ser Ser Gly Ser Leu Trp Ala Ala 690
      695
      700

      Gln Glu Thr Val His Ile Lys Gln Gly Glu Thr Phe Val Tyr Lys Pro 715
      710
      720

 5
       Val Val Glu Gly Pro Ile Pro Asp Gly Ser Tyr Arg Ala Thr Leu His
725 730 735
       Ala Phe Val Asn Gly Gln Gln Gln Leu Tyr Leu Lys Gly Lys Arg Asn 740 750
       Tyr Thr Val Lys lle Val Asn Gly Thr Ala Val Glu Ala Ile Glu Ser
755 760 765
10
       Ser Glu Glu Ile Arg Val Phe Pro Asn Pro Ala Arg Asp Tyr Val Glu
770 775 780
       Ile Ser Ala Pro Cys Ile Pro Gln Glu Thr Ser Ile Ile Leu Phe Asp
785 790 795 800
15
       Leu Ser Gly Lys Ile Val Met Lys Asn Ser Leu Ser Ala Gly His Gly
805 810 815
       Arg Het Asp Val Ser Arg Leu Pro Asn Gly Ala Tyr Ile Leu Lys Val
820 825 830
20
        Asp Gly Tyr Thr Thr Lys Ile Asn Ile Val His
        (2) INFORMATION FOR SEQ ID NO:433
25
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 290 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
30
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
35
                    (A) ORGANISM: Porphyromonas gingivalis
                    (A) HAME/KEY: misc feature
                    (B) LOCATION 1...290
40
             (mi) SEQUENCE DESCRIPTION: SEQ ID NO:433
       45
        Val Ala Ser Cys Asp Ile Ile Asp Lys Asp Gln Thr Leu Leu Pro Ala 20 25 30
        Pro Thr Asn Val Thr Pro Asp Asn Pro Asp Asp Asn Pro Ser Glu Ile 35 40 45
        Asp Ile Thr Gln Thr His Thr Glu Lys Tyr Val Leu Ala Glu Glu Phe 50 55 60
50
        Thr Gly Gln Lys Cys Leu Asn Cys Pro Lys Gly His Arg Lys Leu Ala 65 70 75 80
       Ala Leu Lys Glu Gln Tyr Gly Lys Arg Leu Thr Val Val Gly Ile His
85 90 95
       Ala Gly Pro Gly Ser Leu Val Pro Pro Leu Phe Arg Thr Glu Ala Gly 100 105 110
55
        Asp Ala Tyr Tyr Ser Lys Phe Ala Asn Asn Thr Pro Leu Pro Ala Leu
115 120 125
        Net Val Ser Arg Lys Lys Phe Gly Ser Ser Tyr Val Tyr Asp Lys Ser
130 135
60
       Tyr Lys Thr Trp Asp Val Pro Ile Ala Glu Gln Net Glu Gln Lys Ala
145 150 155 160
Lys Ile Asn Ile Phe Ala Val Ala Glu Tyr Thr Asp Thr Gln Lys Ile
165 170 175
65
        Lys Val Thr Val Lys Gly Lys Ile Leu Glu Gly Asn Thr Leu Pro Lys 180 \hspace{1cm} 185 \hspace{1cm} 190
        Ser Met Val Gln Val Tyr Leu Leu Glu Asp Lys Leu Ile Ala Pro Gln
195 200 205
        Val Asp Gly Asn Thr Thr Val Glu Asn Tyr Glu His Asn His Val Leu 210 215 220

Arg Gly Ala Val Asn Gly Ile Trp Gly Glu Glu Phe Val Asn Leu Lys 225 230 230 235 240
70
        Asp Tyr Leu Tyr Thr Tyr Ala Val Glu Pro Leu Ser Gly Het Ser Phe
                                                   250
        Val Ala Glu Asn Tyr Ser Ile Val Ala Phe Val Tyr Asp Val Gln Thr
75
```

```
260
                                        265
                                                             270
       Phe Glu Val Tyr Asp Val Val His Val Lys Ile Asn Pro Gln Ser Asp
       (2) INFORMATION FOR SEQ ID NO: 434
            (i) SEQUENCE CHARACTERISTICS:
10
                 (A) LENGTH: 223 amino acids
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
15
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
20
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...223
25
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:434
      Phe Ala Gly Cys Gly Leu Asn Asn Met Ala Lys Gly Gly Leu Ile Gly
30
       Ala Gly Val Gly Gly Ala Ile Gly Ala Gly Val Gly Asn Val Ala Gly
35 40
      Asn Thr Ala Val Gly Ala Ile Val Gly Thr Ala Val Gly Gly Ala Ala 50 60
35
       Gly Ala Leu Ile Gly Lys Lys Met Asp Lys Gln Lys Lys Glu Leu Glu
65 70 80
      Ala Ala Val Pro Asp Ala Thr Ile Gln Thr Val Asn Asp Gly Glu Ala
85 90 95
      Ile Leu Val Thr Phe Asp Ser Gly Ile Leu Phe Ala Thr Asn Ser Ser
100 105 110
40
      Thr Leu Ser Pro Asn Ser Arg Thr Ala Leu Thr Lys Phe Ala Ala Asn
115 120 125
      Het Asn Lys Asn Pro Asp Thr Asp Ile Arg Ile Val Gly His Thr Asp 130 135 140
      Asn Thr Gly Ser Asp Lys Ile Asn Asp Pro Leu Ser Glu Arg Arg Ala
145 150 155 160
45
      Ala Ser Val Tyr Ser Phe Leu Asn Ser Gln Gly Val Ser Het Ser Arg
165 170
      Het Ala Ala Glu Gly Arg Gly Ser His Glu Pro Val Ala Asp Asn Ser
180 185 190
50
      Thr Val Ala Gly Arg Ser Ala Asn Arg Arg Val Glu Val Tyr Ile Leu
195 200 205
                                 200
                                                       205
      Pro Asn Ala Lys Met Ile Glu Gln Ala Gln Gln Gly Thr Leu Lys
                               215
55
       (2) INFORMATION FOR SEQ ID NO:435
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 337 amino acids
60
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
65
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
70
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...3\overline{3}7
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:435
75
```

```
Het Ser Lys bys Ser Ile Leu Leu Cu Cys Cys Ser Leu Cys Phe Ile
1 5 10 15
      Ser Ala Thr Lys Ala Val Thr Pro Val Arg Asn Val Arg Asn Ser Gln
20 25 30
 5
      Val Asn Ser Lys Ala Lys Thr Glu Arg Thr Lys Pro Ser Asp Ser Val 35 40 45
      Arg Tyr Ile Ser Asn Met Ile Ala Asp Arg Leu Glu Phe Arg Asn Lys
50 55 60
      Ile Ser Ser Glu Lys Glu Val Arg Lys Ala Glu Tyr Glu Asn Arg Leu
65 70 75 80
10
      Ala Het Glu Ala Leu Asn Tyr Pro Ala Ile Asp Leu Tyr Gly Glu Asp
85 96 95
      Ser Trp Ser Glu Tyr Val Asn Pro Phe Val Gly Ala Gly Thr Asp Val
15
      Glu Ile Pro Asn Ser Tyr Asp Ile Asp Cys Ser Ser Phe Val Met Pro 115 120 125
      Val Glu Asp Lys Gln Val Thr Ser Gln Phe Gly Tyr Arg Arg Arg Phe
130 135 140
      Gly Arg Net His Tyr Gly Ile Asp Leu Ser Val Asn Arg Gly Asp Thr
145 150 155 160
20
      Ile Arg Ala Ala Phe Asp Gly Lys Val Arg Val Arg Ser Tyr Glu Ala
165 170 175
      Arg Gly Tyr Gly Tyr Tyr Ile Val Leu Arg His Pro Asm Gly Leu Glu
180 185 190
      Thr Val Tyr Gly His Het Ser Arg Gln Leu Val Asp Glu Asn Gln Ile
195 200 205
25
      Val Arg Ala Gly Gln Pro Ile Gly Leu Gly Gly Ser Thr Gly Arg Ser
210 215 220
      Thr Gly Pro His Leu His Phe Glu Thr Arg Fhe Het Gly Ile Pro Ile
225 230 235 240
30
      Asn Pro Ser Thr Ile Ile Asp Phe Asp Asn Gly Val Pro Leu Arg Asp
245 250 255
      Ile Tyr Thr Phe Lys Arg Gly Ser Asn Ser Arg Tyr Ala Lys Ala Ser
260 265 270
35
      Lys Thr Ser Ser Arg Tyr Ala Lys Lys Gly Lys Lys Gly Arg Gln Ala
275 280 285
      Ser Ser Pro Met Thr Tyr Arg Ile Lys Lys Gly Asp Thr Leu Glu Thr 290 295 300
      Ile Ala Lys Arg His Gly Thr Ser Val Gln Lys Leu Cys Ala Thr Asn 305 310 315 320
40
      Gly Ile Gly Lys Ser Lys Ile Leu Thr Pro Gly Lys Ala Leu Arg Ile
45
       (2) INFORMATION FOR SEQ ID NO:436
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 151 amino acids
50
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
55
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
60
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...151
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:436
65
      Arg Arg Arg Ile Val Gly Lys Trp Ile Ala Glu Val Cys Ser Arg Tyr
20 25 30
70
      Gly Lys Ala Val Gly Glu Ile Ser Tyr Leu Phe Cys Asp Asp Glu Tyr 35 \, 40 \,
       lle Leu Lys Ala Asn Gl<br/>n Glu Phe Leu Asp His Asp Tyr Tyr Thr Asp 50 \hspace{1.5cm} 55 \hspace{1.5cm} 60 \hspace{1.5cm}
       Ile Ile Thr Phe Asp Ser Cys Glu Ala Asp Thr Val Asp Gly Asp Leu 65 70 75 80
75
```

```
Leu lle Ser Leu Asp Thr Val Arg Ser Asn Ala Arg Ala Leu Asp Leu 85 90 95
        Arg Tyr Glu Asp Glu Leu His Arg Val Ile His Gly Ile Leu His
        Leu Cys Gly Leu Lys Asp Lys Ser Lys Lys Asp Glu Ala Gln Met Arg
115 120 125
        Ala Ala Glu Glu Lys Ala Leu Val Met Leu Arg Glu Thr Ile Gly Ser
        Glu Leu Ser Leu Leu His Thr
10
        (2) INFORMATION FOR SEQ ID NO:437
              (i) SEQUENCE CHARACTERISTICS:
15
                    (A) LENGTH: 391 amino acids (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) NOLECULE TYPE: protein
20
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
25
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...391
30
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:437
        Het Lys Val Lys Tyr Leu Met Leu Thr Leu Val Gly Ala Ile Ala Leu
                                                   10
        Asn Ala Ser Ala Gln Glu Asn Thr Val Pro Ala Thr Gly Gln Leu Pro
20 25 30
35
        Ala Lys Asn Val Ala Phe Ala Arg Asn Lys Ala Gly Ser Asn Trp Phe 35 \hspace{1cm} 40 \hspace{1cm} 45
        Val Thr Leu Gln Gly Gly Val Ala Ala Gln Phe Leu Asn Asp Asn
50 55 60
40
        Asn Lys Asp Leu Met Asp Arg Leu Gly Ala Ile Gly Ser Leu Ser Val
65 75 80
        Gly Lys Tyr His Ser Pro Phe Phe Ala Thr Arg Leu Gln Ile Asn Gly
85 90
        Gly Gln Ala His Thr Phe Leu Gly Lys Asn Gly Glu Gln Glu Ile Asn
100 105 110
45
        Thr Asn Phe Gly Ala Ala His Phe Asp Phe Net Phe Asp Val Val Asn 115 120 125
        Tyr Phe Ala Pro Tyr Arg Glu Asn Arg Phe Phe His Leu Ile Pro Trp
130 135 140
50
        Val Gly Val Gly Tyr Gln His Lys Phe Ile Gly Ser Glu Trp Ser Lys
145 150 160
       Asp Asn Val Glu Ser Leu Thr Ala Asn Val Gly Val Met Met Ala Phe
165 170 175

Arg Leu Gly Lys Arg Val Asp Phe Val Ile Glu Ala Gln Ala Ala His
180 185 190
55
       Ser Asn Leu Ser Arg Ala Tyr Asn Ala Lys Lys Thr Pro Val

195 - 200 - 205

Phe Glu Asp Pro Ala Gly Arg Tyr Tyr Asn Gly Phe Gln Gly Met Ala

210 - 215 - 225
60
       Thr Ala Gly Leu Asn Phe Arg Leu Gly Ala Val Gly Phe Asn Ala Ile
225 230 235 240
        Unk Pro Met Asp Tyr Ala Leu Ile Asn Asp Leu Asn Gly Gln Ile Asn
245 250 255
       Arg Leu Arg Ser Glu Val Glu Glu Leu Ser Lys Arg Pro Val Ser Cys
265 270
65
        Pro Glu Cys Pro Glu Val Thr Pro Val Thr Lys Thr Glu Asn Ile Leu
275 280 285
       Thr Glu Lys Ala Val Leu Phe Arg Phe Asp Ser His Val Val Asp Lys
290 295 300
       Asp Gln Leu Ile Asn Leu Tyr Asp Vai Ala Gln Phe Val Lys Glu Thr
305 310 320
70
       Asn Glu Pro Ile Thr Val Val Gly Tyr Ala Asp Pro Thr Gly Asn Thr 325 330 335
        Gln Tyr Asn Glu Lys Leu Ser Glu Arg Arg Ala Lys Ala Val Val Asp
75
```

```
Val Leu Thr Gly Lys Tyr Gly Val Pro Ser Glu Leu Ile Ser Val Glu
             355
                                     360
                                                         365
       Trp Lys Gly Asp Ser Thr Gln Pro Phe Ser Lys Lys Ala Trp Asn Arg
                              375
 5
       Val Val Ile Val Arg Ser Lys
       (2) INFORMATION FOR SEQ ID NO:438
10
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 385 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
15
           (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
20
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...385
25
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:438
       Het Thr Tyr Arg Ile Met Lys Ala Lys Ser Leu Leu Leu Ala Leu Ala
30
      Gly Leu Ala Cys Thr Phe Ser Ala Thr Ala Gln Glu Ala Thr Thr Gln
20 25 30
      Asn Lys Ala Gly Net His Thr Ala Phe Gln Arg Asp Lys Ala Ser Asp 35 40 45
      His Trp Phe Ile Asp Ile Ala Gly Gly Ala Gly Met Ala Leu Ser Gly 50 55 60
35
      Trp Asn Asn Asp Val Asp Phe Val Asp Arg Leu Ser Ile Val Pro Thr 65, 70 75 80
      Phe Gly Ile Gly Lys Trp His Glu Pro Tyr Phe Gly Thr Arg Leu Gln 85 90 95
       Phe Thr Gly Phe Asp Ile Tyr Gly Phe Pro Gln Gly Ser Lys Glu Arg
40
      Asn His Asn Tyr Phe Gly Asn Ala His Leu Asp Phe Met Phe Asp Leu
115 120 125
       Thr Asn Tyr Phe Gly Val Tyr Arg Pro Asn Arg Val Phe His Ile Ile
130 135 140
45
       Pro Trp Ala Gly Ile Gly Phe Gly Tyr Lys Phe His Ser Glu Asn Ala
145 150 155 160
      Asn Gly Glu Lys Val Gly Ser Lys Asp Asp Het Thr Gly Thr Val Asn
165 170 175
50
      Val Gl; Leu Het Leu Lys Phe Arg Leu Ser Arg Val Val Asp Phe Asn
180 185 190
       Ile Glu Gly Gln Ala Phe Ala Gly Lys Met Asn Phe Ile Gly Thr Lys
195 200 205
      Arg Gly Lys Ala Asp Phe Pro Val Met Ala Thr Ala Gly Leu Thr Phe 210 215 220
55
      Asn Leu Gly Lys Thr Glu Trp Thr Glu Ile Val Pro Met Asp Tyr Ala
225 230 235 240
      Leu Val Asn Asp Leu Asn Asn Gln Ile Asn Ser Leu Arg Gly Gln Val 245 \, 250 \, 255 \,
60
      Glu Glu Leu Ser Arg Arg Pro Val Ser Cys Pro Glu Cys Pro Glu Pro 260 265 270
      Thr Gin Pro Thr Val Thr Arg Val Val Val Asp Asn Val Val Tyr Phe 275 280 285
      Arg Ile Asn Ser Ala Lys Ile Asp Arg Asn Gln Glu Ile Asn Val Tyr
290 295 300
65
      Asn Thr Ala Glu Tyr Ala Lys Thr Asn Asn Ala Pro Ile Lys Val Val 305 310 315
      Gly Tyr Ala Asp Glu Lys Thr Gly Thr Ala Ala Tyr Asn Het Lys Leu
325 330 335
70
      Ser Glu Arg Arg Ala Lys Ala Val Ala Lys Het Leu Glu Lys Tyr Gly 340 345 350
      Val Ser Ala Asp Arg Ile Thr Ile Glu Trp Lys Gly Ser Ser Glu Gln
355 360 365
       Ile Tyr Glu Glu Asn Ala Trp Asn Arg Ile Val Val Met Thr Ala Ala
75
```

Glu 385 (2) INFORMATION FOR SEQ ID NO: 439 5 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 190 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear 10 (ii) MOLECULE TYPE: protein (iii) HYPOTHETICAL: YES 15 (vi) ORIGINAL SOURCE: (A) ORGANISH: Porphyromonas gingivalis (ix) FEATURE: (A) NAME/KEY: misc feature 20 (B) LOCATION 1...190 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:439 Het Glu Phe Phe Het Leu Phe Ile Ala Ala Val Phe Val Asn Asn Val 25 Val Leu Ser Gln Phe Leu Gly Ile Cys Pro Phe Leu Gly Val Ser Lys 20 Lys Val Asp Thr Ser Ile Gly Met Gly Ala Ala Val Thr Phe Val Leu 35 40 45 30 Ala Leu Ala Thr Leu Val Thr Phe Leu Ile Gln L;s Phe Val Leu Asp 55 Arg Phe Gly Leu Gly Phe Met Gln Thr Ile Ala Phe Ile Leu Val Ile 65 70 75 80 Ala Ala Leu Val Gln Met Val Glu Ile Ile Leu Lys Lys Val Ser Pro 35 85 90 Pro Leu Tyr Gln Ala Leu Gly Val Phe Leu Pro Leu Ile Thr Thr Asn 100 105 110 Cys Cys Val Leu Gly Val Ala Ile Leu Val Ile Gln Lys Asp Tyr Thr 120 125 40 Leu Leu Gln Ser Phe Val Tyr Ala Ile Ser Thr Ala Ile Gly Phe Thr 130 135 140 Leu Ala Met Val Thr Phe Ala Gly Ile Arg Glu Gln Leu Asp Met Thr 145 150 155 Asn Leu Pro Lys Ala Met Lys Gly Ile Pro Ser Ala Leu Leu Ala Ala 165 170 175 45 Gly Ile Leu Ala Het Ala Phe Met Gly Phe Ser Gly Ile Ala 180 195 (2) INFORMATION FOR SEQ ID NO: 440 50 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 186 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear 55 (ii) MOLECULE TYPE: protein (iii) HYPOTHETICAL: YES 60 (vi) ORIGINAL SOURCE: (A) ORGANISH: Porphyromonas gingivalis (ix) FEATURE: (A) NAME/KEY: misc_feature 65 (B) LOCATION 1...186 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:440 Het Leu Phe Ile Ala Ala Val Phe Val Asn Asn Val Val Leu Ser Gln 70 10 Phe Leu Gly Ile Cys Pro Phe Leu Gly Val Ser Lys Lys Val Asp Thr 25 30 Ser Ile Gly Het Gly Ala Ala Val Thr Phe Val Leu Ala Leu Ala Thr 40 75 Leu Val Thr Phe Leu Ile Gln Lys Phe Val Leu Asp Arg Phe Gly Leu

```
Gly Phe Het Gln Thr Ile Ala Phe Ile Leu Val Ile Ala Ala Leu Val
65 70 80
       Gln Het Val Glu Ile Ile Leu Lys Lys Val Ser Pro Pro Leu Tyr Gln
85 90
 5
       Ala Leu Gly Val Phe Leu Pro Leu Ile Thr Thr Asn Cys Cys Val Leu
100 105 110
       Gly Val Ala Ile Leu Val Ile Gln Lys Asp Tyr Thr Leu Leu Gln Ser
115 120 125
10
       Phe Val Tyr Ala Ile Ser Thr Ala Ile Gly Phe Thr Leu Ala Met Val
130 135 140
       Thr Phe Ala Gly Ile Arg Glu Gln Leu Asp Met Thr Asn Leu Pro Lys
145 150 155 160
       Ala Met Lys Gly Ile Pro Ser Ala Leu Leu Ala Ala Gly Ile Leu Ala
165 170 175
15
       Net Ala Phe Net Gly Phe Ser Gly Ile Ala
       (2) INFORMATION FOR SEQ ID NO:441
20
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 833 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
25
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
30
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
35
                  (B) LOCATION 1...833
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:441
      Met Lys Gln Leu Asn Ile Ile Ser Phe Ile Ile Ala Phe Leu Phe Leu 1 5 10 15
40
       Gly Thr Ser Ala Ser Ala Gln Gln Ser Gly Gly Ser Val Thr Gly Thr
20 25 30
       Val Val Asp Lys Ser Ser Lys Glu Pro Ile Ala Tyr Val Gln Val Phe
35 40 45
       Val Lys Gly Thr Thr Leu Gly Thr Ser Thr Asp Ala Asn Gly Asn Tyr 50 60
45
       Ser Ile Lys Gly Ile Pro Ser Gly Asn Gln Thr Ile Val Ala Arg Leu
65 70 80
       Het Gly Tyr Ser Thr Cys Glu Glu Lys Val His Ile Glu Lys Gly Gly
85 90 95
50
       Ser Arg His Val Asp Leu Tyr Leu Thr Glu Glu Ile Leu Ser Leu Asp
100 105 110
       Gly Val Val Val Ser Ala Asn Arg Asn Glu Thr Phe Arg Arg Gln Ala 115 120 125
       Pro Ser Leu Val Thr Val Leu Ser Pro Glu Leu Phe Leu Lys Thr Asn
130 135 140
55
       Ser Thr Asn Leu Ser Gln Gly Leu Lys Phe Gln Pro Gly Leu Arg Val 145 150 155 160
       Glu Asp Asn Cys Gln Asn Cys Gly Phe Asn Gln Val Arg Ile Asn Gly
165 170 175
60
       Leu Glu Gly Ala Tyr Ser Gln Ile Leu Ile Asp Ser His Pro Ile Phe 180\,
       Ser Ser Leu Ala Gly Val Tyr Gly Leu Glu Gln Met Pro Ala Asn Met 195 200 205
65
       Ile Glu Arg Val Glu Val Ile Arg Gly Gly Gly Ser Ala Leu Phe Gly 210 215 220
       Ser Asn Ala Val Gly Val Ile Asn Val Ile Thr Lys Glu Pro Leu
225 230 235 249
       Arg Asn Ser Ala Glu Ile Ser His Ser Thr Met Thr Phe Asp His Ala
245 250 255
70
       Lys Gly Trp Gly Ser Phe Gln Asn Thr Thr Gln Phe Asn Gly Ser Met 260 265 270
       Leu Thr Glu Asp Arg Lys Ala Gly Val Het Val Phe Gly Gln His Asn 275 280 285
75
       Tyr Arg Pro Gly Gln Asp Ile Asp Gly Asp Asn Phe Thr Glu Leu Pro
```

```
295
       Asn Leu Arg Asn Arg Ser Leu Gly Phe Arg Ser Tyr Tyr Lys Thr Gly 305 310 315
       Leu Tyr Ser Lys Ala Thr Leu Glu Tyr His Ser Met Gln Glu Tyr Arg
325 330 335
 5
       Arg Gly Gly Asp Arg Leu Asp Asn Pro Pro Phe Glu Ala Gln Ile Ala 340 345 350
       Glu Tyr Leu Gln His Tyr Ile Asn Gly Gly Ser Phe Lys Phe Asp Gln 355 360 365
10
       Gly Phe Ser Gly Gly Lys Asp Phe Phe Ser Leu Tyr Ala Ser Ala Gln 370 375 380
       Asp Val Gln Arg Arg Ser Tyr Tyr Gly Gly Gly Asp Tyr Thr Glu Asn
385 390 395 400
       Leu Leu Asn Gly Ala Val Gln Ser Gly Ser Thr Glu Ser Asp Glu Tyr
405 410 415
15
       Asn Asp Ala Phe Thr Ala Leu Thr Ser Tyr Gly Thr Thr Lys Gly Phe 420 430
       Asp Leu Gln Gly Gly Gly Het Tyr Arg His Thr Phe Gly Glu Asn Trp 435 440 445
20
       Asp Phe Thr Gly Gly Leu Glu Tyr Ile Tyr Gly Gln Leu Asp Asp Arg 450 460
       Ser Gly Tyr Arg Pro Ser Lys Ile Asp Gln Asn Thr Ser Thr Phe Ser 465 -470 -475 -480
       Gln Tyr Asp Gln Leu Glu Tyr Lys Thr Glu Lys Leu Ser Ala Leu Ile
485 490 495
25
       Gly Ala Arg Ile Asp Tyr Val Leu Leu Asn Gln Asp Gly Lys Arg Tyr 500 505 510
       Ile Asp Pro Leu Phe Ile Phe Ser Pro Arg Ala Asn Val Arg Tyr Asn
515 520 525
30
       Pro Asn Lys Asn Leu Ser Phe Arg Leu Ser Tyr Ser Glu Gly Phe Arg
530 540
       Ala Pro Gln Tyr Phe Asp Glu Asp Leu His Val Glu Leu Ala Gly Gly 545 550 555 560
       Thr Pro Ile Ser Arg Val Leu Ser Pro Asn Leu Lys Glu Glu Arg Ser
565 570 575
35
       Arg Ser Ile Ser Ala Ser Fhe Asp Tyr Tyr His Arg Ala Asp Glu Trp
580 585 590
       Gln Phe Asn Ile Met Gly Glu Ala Phe Ser Thr Phe Ile Ser Asn Gln 595 600 605
       Phe Lys Pro Ser Asp Lys Val Glu Thr Thr Ser Asp Gly Lys Glu Trp 610 615 620

Ile Ile Arg Thr Ile Tyr Asn Asp Lys Asp Gl; Val Ser Lys Val Tyr 625 630 635 640
40
       Gly Val Asn Leu Glu Gly Arg Ile Ala Tyr Asn Lys Ser Phe Asp Leu
645 650 655
45
       Gln Leu Gly Gl; Thr Trp Gln Arg Ser Arg Tyr Gly Ser Ile Tyr Thr
660 665 670
       Ala Val Glu Ala Asp Lys Thr Thr Gly Gln Ala Glu Ile Ser Val Lys 675 680 685
                                       680
50
       Asp Tyr Val Arg Thr Pro Asn Leu Tyr Gly Tyr Phe Val Ala Thr Val 690 700
       Arg Pro Thr Glu His Phe Ala Ile Asn Leu Ser Gly Thr Phe Thr Gly 705 710 715 720
       Lys Het Asp Val Val His Glu Ala Tyr Glu Giy Asp Ile Pro Ala Glu
725 730 735
55
       His Ile Ala Pro Asp Gly Ser Phe Asp Phe Glu Het Asn Gly Gln Gln 740 750
       Phe Lys Gly Leu Ala Glu Gly His Ala Lys Leu Val Lys Thr Pro Ala 755 760 765
60
       Phe Ala Asp Ile Asp Leu Lys Leu Ser His Asp Phe His Leu Ala Ser
770 780
       Thr Het Thr Leu Glu Leu Asn Ala Gly Ile Gln Asn Ile Phe Asn Ser
785 790 800
       Tyr Gln Lys Asp Thr Asp Lys Gly Pro Gly Arg Ala Ser Thr Tyr Val
805 810 815
65
       Tyr Gly Pro Met Gln Pro Arg Arg Ile Phe Val Gly Thr Lys Ile Asn
70
        (2) INFORMATION FOR SEQ ID NO: 442
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 891 amino acids
```

(B) TYPE: amino acid

75

```
(D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
 5
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
10
            (ix) FEATURE:
                   (A) HAME/KEY: misc feature
                   (B) LOCATION 1...891
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:442
15
       Het Tyr Lys Lys Ile Ile Ala Val Ala Ala Leu Phe Cys Ala Ser Ile
       Gly Ile Leu Lys Gly Gln Ser Ser Asp Leu Thr Pro Gln Asp Thr Ile 20 25 30
       Tyr Ser Pro Glu Ile Ser Tyr Ala Lys Pro Ile His Lys Thr Ile Ala 35 40 45
20
       Ser Ile Glu Ile Glu Gly Het Arg Ser Phe Asp Asp Phe Val Leu Arg 50 60
       Asn Leu Ser Gly Leu Ala Val Gly Asp Glu Val Leu Ile Pro Gly Asp 65 70 80
25
       Ala Het Ser Ala Ala Val Asm Arg Ile Het Arg Glm Gly Tyr Phe Ser
85 90 95
       Asn Val Arg Ile Ile Ala Asp Lys Tyr Val Gly Asn Lys Val Tyr Leu 100 105 110
30
       Lys Ile Ile Val Thr Glu Arg Pro Arg Ile Ser Lys Val Thr Phe Ser
115 120 125
       Gly Val Lys Lys Ser Glu Arg Glu Asp Leu Glu Met Lys Ile Gly Leu
130 135 140
       Arg Glu Gly Ile Gln Met Thr Arg Asn Asn Glu Asp Lys Val Arg Gln
145 150 155 160
35
       The Val Gln Lys Tyr Phe Ser Glu Lys Gly Tyr Arg Asp Ala Ser Ile
165 170 175
       Arg Ile Thr Gln Glu Pro Asp Leu Ser Lys Asp Gly Phe Val Asn Val
40
       Leu Ile Ser Ile Glu Lys Lys Ser Lys Thr Lys Val Asn Glu Ile Tyr
195 200 205
        Phe Ser Gly Asn Lys Ala Leu Ser Asn His Lys Leu Arg Met Ala Met
210 215 220
       Lys Asn Thr Asn Ala Lys Phe Ser Leu Arg Lys His Ile Arg Ser Ser 225 230 235 240
45
       Phe Leu Lys Leu Phe Ser Thr His Lys Phe Val Glu Glu Ser Tyr Arg
245 250 255
       Glu Asp Leu Val Arg Leu Ile Glu Lys Tyr Gln Glu Tyr Gly Tyr Arg 260 265 270

Asp Ala Glu Ile Leu Thr Asp Ser Val Val Lys Ala Pro Asp Gly Lys 275 280 285
50
       Arg Val Asp Ile Tyr Leu Asn Ile Glu Glu Gly Gln Lys Tyr Tyr Ile
290 295 300
       Lys Asp Val Asn Phe Val Gly Asn Ser Gln Tyr Pro Ser Glu Tyr Leu
305 310 315 320
55
       Glu Arg Val Leu Gly Ile Lys Ser Gly Asp Val Tyr Asn Gln Arg Arg
325 330 335
       Leu Ala Lys Arg Leu Asn Glu Asp Glu Asp Ala Val Gly Asn Leu Tyr 340 345 350
60
       Tyr Asn Asn Gly Tyr Ile Phe Ala Trp Val Asp Pro Val Glu Thr Asn 355 360 365
        Val Val Gly Asp Ser Val Ser Leu Asp Ile Arg Ile Ala Glu Gly Lys
370 375 380
       Gin Ala Asn Ile Asn Lys Val Ile Ile Lys Gly Asn Thr Val Val Tyr
385 390 395 460
65
       Glu Asp Val Val Arg Arg Glu Leu Tyr Thr Lys Pro Gly Gln Leu Phe
405 410 415
        Ser Arg Glu Asp Ile Ile Asn Ser Ile Arg Leu Ile Asn Gln Leu Gly 420 425 430
70
        His Phe Asp Ala Glu Lys Ser Ile Pro Arg Pro Ile Pro Asn Pro Glu
435 440 445
        Thr Gly Thr Val Asp Ile Glu Tyr Asp Leu Val Pro Arg Ser Ser Asp
450 455 460
        Gln Leu Glu Leu Ser Val Gly Trp Ser Gln Ser Gly Leu Leu Phe Arg
465 470 475
75
```

```
Gly Ala Ile Lys Phe Thr Asn Phe Ser Val Gly Asn Leu Leu His Pro 485 490 495
      Ser Net Tyr Lys Lys Gly Ile Ile Pro Gln Gly Asp Gly Gln Thr Leu 500 505 510
      Ser Leu Ser Ala Gln Thr Asn Gly Lys Tyr Tyr Gln Gln Tyr Ser Val
515 520 525
      Thr Phe Met Asp Pro Trp Phe Gly Gly Lys Arg Pro Asp Net Phe Ser 530 540
      Phe Ser Ala Phe Tyr Ser Lys Thr Thr Ala Ile Asp Ser Lys Phe Tyr 545 550 560
10
      Asn Ser Asn Ala Gly Asn Tyr Tyr Asn Ala Tyr Tyr Asn Ser Tyr Tyr
565 570
      Asn Asn Tyr Asn Ser Tyr Tyr Asn Gly Met Ser Asn Tyr Thr Gly Asp
580 585 590
      Leu Tyr Thr Gln Ala Ser Asp Pro Asp Arg Ser Leu Gln Met Leu Gly 595 600 605
15
      Thr Ser Ile Gly Tyr Gly Lys Arg Leu Thr Trp Pro Asp Asn Trp Phe 610 615 620
      Gln Ile Tyr Thr Ser Leu Asn Tyr Thr Tyr Tyr Arg Leu Arg Asn Trp 625 630 635 640
20
      25
      Thr Arg Ser Gly Ser Asp Phe Het Val Ser Val Ala Ala Thr Leu Pro
675 680 685
      Tyr Ser Leu Trp Asp Asn His Asp Tyr Ala Ser Gln Asn Leu Ser Val690 \hspace{1cm} 695 \hspace{1cm} 700
      Ser Asp Arg Tyr Arg Phe Ile Glu Tyr His Lys Trp Lys Phe Arg Gly 705 710 715 720
30
      Arg Val Phe Thr Pro Leu Leu Asn Pro Ala Thr His Lys Tyr Thr Pro 725 730 735
      Val Leu Het Ser Arg Val Glu Gly Ala Val Leu Gly Ser Tyr Asn Ser
740 745 750
35
      Asn Lys Lys Ser Pro Phe Gly Thr Phe Tyr Met Gly Gly Asp Gly Met 755 760 765
      Ser Ser Tyr Tyr Gly Gly Tyr Met Asn Glu Thr Ile Gly Leu Arg Gly 770 780
      Tyr Lys Asn Gly Ser Ile Ala Gly Asn Asn Tyr Asp Tyr Ala Tyr Ala
785 790 795 800
40
      Tyr Net Arg Leu Thr Net Glu Leu Arg Phe Pro Ile Leu Phe Glu Asn 805 810 815
      Ser Phe Asn Ala Trp Leu Leu Ala Phe Ala Glu Ala Gly Asn Ala Trp
820 825 830
      Arg Ser Ile Asp Asn Tyr Asn Pro Phe Asn Leu Lys Arg Ser Ala Gly 835 840
45
      Val Gly Leu Arg Val Thr Leu Pro Met Val Gly Het Leu Gly Ile Asp
850 855 860
      Trp Gly Tyr Gly Phe Asp Arg Pro Asp Asn Ser Leu Gln Arg Gly Gly
                         870
                                                875
      Ser Asn Val His The Val Leu Gly Gln Glu Phe
                        885
       (2) INFORMATION FOR SEQ ID NO:443
55
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 174 amino acids
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
60
           (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
65
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) MAME/KEY: misc_feature
70
                 (B) LOCATION 1...174
           (Mi) SEQUENCE DESCRIPTION: SEQ ID NO:443
      Met Asn Gly Asp Met Lys Arg Phe Leu Ile Leu Ile Gly Phe Ala Leu
75
```

```
Ala Val Ala Phe Ser Gly Phe Ser Gln Lys Phe Ala Leu Val Asp Met
                                         25
                   20
      Glu Tyr Ile Leu Arg Asn Ile Pro Asp Tyr Glu Met Met Asn Glu Gln
                                   40
      Leu Glu Gln Val Ser Lys Lys Trp Gln Asn Glu Ile Glu Ala Leu Glu
50 55 60
 5
      Asn Glu Ala Gln Ser Met Tyr Lys Lys Tyr Gln Ser Asp Leu Val Phe 65 70 75 80
      Leu Ser Ala Ala Gln Lys Lys Thr Gln Glu Glu Ala Ile Val Lys
85 90 95
10
      Glu Gln Gln Ala Ser Glu Leu Lys Arg Lys Tyr Phe Gly Pro Glu Gly
100 105 110
      Glu Leu Tyr Lys Lys Arg Ser Asp Leu Het Lys Pro Ile Gln Asp Glu
115 120 125
      Ile Trp Asn Ala Ile Lys Glu Ile Ala Lys Arg Asn Asn Tyr Gln Het
130 135 140
15
      Val Leu Asp Arg Gly Thr Ser Gly Ile Ile Phe Ala Ser Pro Ser Ile
                                               155
                            150
      Asp Ile Ser Asp Leu Val Leu Ser Lys Het Gly Phe Ser Lys
20
                        165
       (2) INFORMATION FOR SEQ ID NO: 444
            (i) SEQUENCE CHARACTERISTICS:
25
                  (A) LENGTH: 170 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
30
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
35
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
(B) LOCATION 1...170
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:444
       Met Lys Arg Phe Leu Ile Leu Ile Gly Phe Ala Leu Ala Val Ala Phe
                                              1.0
       Ser Gly Phe Ser Gln Lys Phe Ala Leu Val Asp Net Glu Tyr Ile Leu
20 25 30
45
                  20
       Arg Asn Ile Pro Asp Tyr Glu Met Met Asn Glu Gln Leu Glu Gln Val 35 40 45
       Ser Lys Trp Gln Asn Glu Île Glu Aia Leu Glu Asn Glu Ala Gln
50 55 60
       Ser Met Tyr Lys Lys Tyr Gln Ser Asp Leu Val Phe Leu Ser Ala Ala 65 70 80
50
       Gln Lys Lys Thr Gln Glu Glu Ala Ile Val Lys Lys Glu Gln Gln Ala
85 90 95
       Ser Glu Leu Lys Arg Lys Tyr Phe Gly Pro Glu Gly Glu Leu Tyr Lys
100 105 110
55
       Lys Arg Ser Asp Leu Met Lys Pro Ile Gln Asp Glu Ile Trp Asn Ala
115 120 125
       Ile Lys Glu Ile Ala Lys Arg Asn Asn Tyr Gln Met Val Leu Asp Arg
130 135 140
       Gly Thr Ser Gly Ile Ile Phe Ala Ser Pro Ser Ile Asp Ile Ser Asp
145 150 155
60
       Leu Val Leu Ser Lys Met Gly Phe Ser Lys
 65
        (2) INFORMATION FOR SEQ ID NO:445
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 163 amino acids
                   (B) TYPE: amino acid
 70
                   (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
 75
```

```
(vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
 5
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...163
           (mi) SEQUENCE DESCRIPTION: SEQ ID NO:445
10
      Het Lys Lys Phe Phe Leu Met Leu Leu Met Ala Leu Pro Leu Ser Leu
                                              1.0
      Leu Ala Gl<br/>n Lys Val Ala Val Val Asn Thr Glu Glu Ile Ile Ser Lys<br/> 20 25 30
      Het Pro Glu Gln Val Ala Ala Thr Lys Gln Leu Asn Glu Leu Ala Glu
35 40 45
15
      Lys Tyr Arg Leu Asp Leu Lys Ser Het Asp Asp Glu Phe Ala Lys Lys 50 60
      Thr Glu Glu Phe Val Lys Glu Lys Asp Ser Leu Leu Glu Asn Ile Arg 65 70 70 80
      Asn Arg Arg Gln Gln Glu Leu Gln Asp Ile Gln Thr Arg Tyr Gln Gln
85 90 95
20
      Ser Tyr Gln Thr Met Gln Glu Asp Leu Gln Lys Arg Gln Gln Gln Leu
100 105 110
      Fhe Ala Pro Ile Gln Gln Lys Val Ala Asp Ala Ile Lys Lys Val Gly 115 120
25
      Asp Glu Glu Asn Cys Ala Tyr Ile Het Glu Ala Gly Het Het Leu Tyr
130 135 140
      Thr Gly Ala Thr Ala Ile Asp Leu Thr Ala Lys Val Lys Ala Lys Leu
30
      Gly Ile Lys
       (2) INFORMATION FOR SEQ ID NO:446
35
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 827 amino acids(B) TYPE: amino acid
                 (D) TOPOLOGY: linear
40
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
45
                 (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...827
50
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:446
      Het Lys Glu Ala Ile Pro Arg Lys Asn Lys Tyr Ile Lys Leu Asn Gly 1 \phantom{\bigg|} 15
55
      Ala Ala Met Ala Gln Gly Val Arg Val Ser Gly Tyr Val Leu Asp Arg
35 40 45
      Gly Glu Lys Pro Ile Pro Phe Ala Gly Val Lys Val Arg Gly Thr Gly 50 60
60
      Thr Gly Ala Thr Thr Asn Leu Lys Gly Tyr Tyr Glu Phe Arg Met Lys 65 70 75 80
      Ala Thr Thr Asp Ser Ile Thr Ile Glu Phe Ser Ser Het Gly Tyr Gln
85 90
      Gly Val Ser Arg Ser Phe Pro Ser Leu Thr Lys Asp Thr Arg Leu Asn
100 105 110
65
      Val Arg Leu Ala Glu Ala Glu Het Glu Leu Ser Ser Val Thr Val Gln
115 120 125
      Ala Thr Lys Arg Arg Leu Asn Thr Net Glu Arg Val Asn Thr Arg Asp
130 135 140
70
      Leu Arg Val Asn Ala Gly Pro Thr Gly Gly Val Glu Ser Leu Ile Ser
145 150 160
      Thr Tyr Ala Gly Val Thr Gln Asn Asn Glu Leu Ser Ser Gln Tyr Ser
165 170 175
75
      Val Arg Gly Gly Ser Tyr Asp Glu Asn Met Val Tyr Val Asn Gly Val
```

| | Glu | Val | | 180 Arg | Pro | Геп | Leu | | 185 Arg | Ser | Ala | Gln | Gln 205 | 190 Glu | Gl y | Leu |
|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|
| _ | Ser | | 195 Val | Asn | Pro | Asp | | 200 Thr | Gln | Ser | Val | Gln 220 | | Ser | Ala | G1 y |
| 5 | Gly 225 | 210 Phe | Thr | Ala | Asp | Tyr 230 | 215 Gly | Asp | Lys | Met | Ser 235 | | Val | Leu | Asp | 11e 240 |
| | Arg | T5-r | Γ?.ε | Gln | Pro 245 | | Glu | Γ.'.ε | Glu | G15. 250 | | Val | Leu | Leu | Gly 255 | |
| 10 | Leu | Gln | Ser | Ser 260 | Ala | Tyr | Ty.r | G17 | Ser 265 | | Ala | G1λ. | Ala | Phe 270 | | Gln |
| | Ile | Thr | G1 y 275 | | Arg | Туr | Lys | Ser 280 | | Lys | Ser | Leu | Leu 285 | | Thr | Thr |
| 15 | Asp | Thr 290 | Lys | Ala | Glu | Туr | Asp 295 | | Ile | Tyr | Ala | Asp 300 | G17 | Gln | Thr | Phe |
| | 305 | Thr | | | Phe | 310 | Pro | | | | 315 | | | | | 320 |
| | lle | Ser | Gln | Thr | Arg 325 | тут | Lys | Phe | Val | Pro 330 | Gln | Thr | Arg | Glu | Thr 335 | Ser |
| 20 | | | | 340 | Ser | | | | 345 | | | | | 350 | | |
| | | | 355 | | Arg | | | 360 | | | | | 365 | | | |
| 25 | | 370 | | | Asp | | 375 | | | | | 380 | | | | |
| | 385 | | | | Arg | 390 | | | | | 395 | | | | | 400 |
| • | | - | | | Leu 405 | | | | | 410 | | | | | 415 | |
| 30 | | | | 420 | Asn | | | | 425 | | | | | 430 | | |
| | | | 435 | | Ser | | | 440 | | | | | 445 | | | |
| 35 | | 450 | | | Glu | | 455 | | | | | 460 | | | | |
| | 465 | | | | Ala | 470 | | | | | 475 | | | | | 480 |
| 4 0 | | | | | Leu 485 | | | | | 490 | | | | | 495 | |
| 40 | | | | 500 | Asp Phe | | | | 505 | | | | | 510 | | |
| | | | 515 | | Arg | | | 520 | | | | | 525 | | | |
| 45 | | 530 | | | Ile | | 535 | | | | | 540 | | | | |
| | 545 | | | | Ala | 550 | | | | | 555 | | | | | 560 |
| 50 | | | | | 565 Gln | | | | | 570 | | | | | 575 | |
| | | | | 580 | | | | | 585 | | | | | 290 | | |
| | | | 595 | | Thr | | | 600 | | | | | 605 | | | |
| 55 | | 610 | | | Тут | | 615 | | | | | 620 | | | | |
| | 625 | | | | Ile | 630 | | | | | 635 | | | | | 640 |
| 60 | | | | | 645 Asp | | | | | 650 | | | | Pro | 655 | |
| | Asp | Ser | Trp | 660 Leu | Thr | Ala | Ser | Ile | 665 Ile | Lys | Ala | Arg | | | Leu | Asp |
| 65 | Gly | | | Ser | Leu | Fro | | | | Ala | Pro | | | | Phe | Ser |
| 65 | | | | Gln | Glu | | | | Gly | Asn | | | | Thr | Ala | |
| | 705 Leu | | Ala | Αla | Leu | 710 Ser | | Glγ | Leu | | | | Asn | Pro | | |
| 70 | Gly | Leu | Ser | | 725 Pro | Ala | Phe | Thr | | | | T;'r | Lys | Arg 750 | | |
| | Leu | Gly | | | Tyr | r?.e | Trp | | | | Asp | Asp | Ser 765 | Phe | | Gly |
| 7 5 | Arg | | | | Leu | Met | | | | Gly. | Ala | T;r | Ile | | Ala | Asp |
| , , | | 770 | | | | | 775 | | | | | 100 | | | | |

```
Leu Phe Asn Leu Phe Asp Met Thr Asn Val Asn Ser Tyr Tyr Trp Val 785 790 795 800
       Ser Asp Ala Tyr Gln Gln Gln Tyr Ala Val Pro Asn Tyr Leu Thr Arg
805 810 815
       Arg Gln Phe Asn Leu Arg Leu Leu Val Glu Phe
        (2) INFORNATION FOR SEQ ID NO:447
10
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 672 amino acids (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
15
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
20
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...672
25
             (mi) SEQUENCE DESCRIPTION: SEQ ID NO:447
       Het Tyr Ser Gly His His Lys Ile His Tyr Pro Phe Leu Ile Leu Leu 1 \phantom{\bigg|}5\phantom{\bigg|} 15
30
       Val Cys Leu Ala Phe Ala Ala Cys Lys Ser Val Lys Leu Lys Asp Ala
20 25 30
       Glu Lys Ala His Asp Arg Gln Glu Tyr Thr Lys Ala Ala Asp Met Tyr 35 40 45
       Asn Thr Leu Tyr Arg Arg Thr Arg Arg Lys Gln Val Glu Met Lys Ala 50 60
35
       Tyr Thr Ala Phe Arg Ser Gly Glu Asn Tyr Arg Ala Ala Gly Arg Gln 65 70 80
       Ala Lys Ala Leu Arg Gly Tyr Leu Asn Ala Arg Arg Tyr Gly Tyr Pro85 \\ 90 \\ 95
40
       Asp Ser Val Val Leu Leu Arg Leu Ala Gin Thr Tyr Gln Gln Gly Gly 100 105 110
       Asn Tyr Lys Glu Ala Glu Val Leu Phe Arg Gly Tyr Leu Glu Ala Tyr
115 120 125
        Pro Lys Ser Tyr Phe Ala Ala Ile Gly Leu Glu Gly Cys Leu Phe Ala
130 135 140
45
       Arg Gln Gln Lys Glu Tyr Pro Thr Arg Tyr Arg Ile Arg Arg Ala Ala
145 155 160
       Glu Trp Asn Ser Ala Arg Gly Asp Phe Gly Pro Ala Tyr Ala Pro Asp

165 176 175

Ala Ser Ala Leu Tyr Phe Thr Ser Ser Arg Ser Lys Asp Asp Gly Leu

180 185 190
50
       Asp Asn Ser Ser Ile Thr Gly Leu Lev Pro Asn Asp Ile Tyr Ile Ile 195 - 200 - 205

Lys Arg Asp Ala Gln Gly Arg Trp Gly Arg Pro Asp Ser Val Ser Gly 210 - 215 - 220
55
       Gly Ile Asn Thr Pro Trp Asp Glu Gly Val Pro Thr Ile Thr Pro Asp 225 230 235 240 Gly Ser Thr Ile Tyr Tyr Thr Leu Ala Gln Gln Gly Ala Asp Tyr Asp 245 250 255
60
       Arg Thr Val Gln Ile Tyr Ser Ala Ala Arg Ser Gly Glu Gly Gly Trp 260 265 270
        Ala His Pro Ser Met Ser Ala Ser Gly Asp Tyr Leu Tyr Phe Val Ser 290 295 300
65
       Asn Ile Gly Gly Ser Tyr Gly Gly Lys Asp Ile Tyr Arg Val Lys Val 305 310 315
       Ser Asp Arg Ser Tyr Gly Ser Pro Glu Asn Leu Gly Pro Asp Ile Asn
325 330 335
70
       Thr Pro Gly Asp Glu Net Phe Pro Phe Ile Asp Gly Asp Ser Thr Leu 340 345 350
        Phe Phe Ala Ser Asp Gly His Ala Gly Leu Gly Gly Leu Asp Ile Phe 355 360 365
        Lys Ala Thr Leu Asp Ser Thr Gly Gln Trp His Val Val Asn Het Gly
75
```

```
Gln Pro Val Asn Ser Ser Ala Asp Asp Phe Gly Leu Ala Val Glu Pro
385 390 395 400
       Lys Gly Lys Asn Lys Glu Glu Ala Leu Pro Asp Asn Gly Val Lys Gly
405 415
       Val Phe Cys Ser Asn Arg Gly Asp Ala Arg Gly Trp Pro His Leu Phe 420 425 430

His Phe Glu Leu Pro Ala Ile Tyr Thr Glu Ile Gln Gly Tyr Val Met 435 445
       Asp Arg Glu Glu Asn Pro Ile Ala Gly Ala Thr Val Arg Ile Val Gly 450 455 460
10
       Glu Arg Gly Pro Val Gly Gln Gly Phe Val Thr Thr Arg Asp Asp Gly
465 470 475 480
       Ser Tyr Lys Met Ser Val Gln Gly Asp Thr Arg Tyr Val Met Leu Ala
485 490 495
       Gly Ala Ser Gly Tyr Leu Asn Gln Tyr Val Glu Leu Lys Thr Asp Thr 500 505 510
15
       Ala Lys Gln Ser Glu Thr Tyr Tyr Val Asp Phe Phe Leu Ala Ser Arg
515 520 525
       Glu Lys Ala Glu Gly Leu Gln Asn Ile Phe Tyr Asp Phe Asp Lys Ala
530 535 540
20
       Thr Leu Arg Pro Glu Ser Met Lys Ser Leu Asp Glu Leu Ile Arg Ile
545 550 555 560
       Leu Thr Asp Asn Pro Asp Ile Arg Ile Glu Leu Gly Ser His Ala Asp
565 575
       Arg Lys Gly Pro Asp Ala Tyr Asn Leu Gly Leu Ser Asp Arg Ala
580 585 590
25
                     580
       Lys Ser Val Val Asp Tyr Leu Thr Ser Arg Gly Ile Ala Ala Asp Arg
595 600 605
       Leu Thr Trp Lys Gly Tyr Gly Lys Ser Val Pro Lys Thr Val Thr Ala
610 620
30
       Lys Ile Ala Glu Arg His Asp Phe Leu Lys Glu Gly Asp Val Leu Thr
625 630 635 640
       Glu Glu Phe Val Ala Pro Leu Thr Glu Glu Gln Gln Ser Val Cys Asp
645 655
       Gln Leu Asn Arg Arg Thr Glu Phe Arg Val Ile Glu Glu Leu Arg
35
        (2) INFORMATION FOR SEQ ID NO:448
40
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 708 amino acids (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
45
             (ii) HOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
50
                   (A) ORGANISH: Perphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                    (B) LOCATION 1...708
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:448
        Het Lys Lys Phe Phe Phe Ala Leu Leu Ser Ile Gly Ile Ser Ala Glu 1 5 15
        Ala Phe Ala Lys Thr Asp Asn Val Pro Thr Asp Ser Leu Arg Val His
60
        Asn Leu Gln Thr Val Thr Val Tyr Ser Thr Arg Thr Ala Val Pro Leu
35 40
        Lys Lys Ile Pro Ala Lys Met Glu Leu Ile Ser Ser Arg Asn Ile Lys 50 \\
65
        Gln Ser Gly Phe Asn Asn Het Thr Asp Ile Leu Lys Thr Gln Ser Ser 65 70 75 80
        Leu Asp Val Ile Gln Tyr Pro Gly Phe Ser Ser Asn Ile Gly Ile Arg
85 90 95
        Gly Phe Lys Pro Ser Gly Lys Tyr Val Thr Val Leu Val Asn Gly Ile
100 105 110
70
        Pro Ala Gly Thr Asp Asn Ile Ser Thr Leu Asn Thr Ser Asn Ile Glu
115
120
125
        Gln Ile Glu Ile Leu Lys Gly Pro Phe Ser Ser Ile Tyr Gly Thr Asn
130 135 140
75
```

| | 145 | | | | | 150 | | | | | 155 | | | | Asp | 160 |
|------------|-------------|------------|------------|------------|------------|------------|-------------|------------|---------------------------|-------------|------------|------------|------------|------------|------------|------------|
| _ | Ile | His | Gly | Asn | Val 165 | Ser | Leu | Phe | G1 y | G1 y 170 | Ser | Tyr | Gln | Thr | Het 175 | Ala |
| 5 | G17. | Ser | Phe | Asn 180 | Leu | Gly | Gly | Arg | Phe 185 | Glu | Asp | Ile | Phe | Ser 190 | Phe | Asp |
| | Leu | Ser | Leu 195 | GJ À. | Leu | Asp | rys | Gln 200 | Aεn | Lys | Asp | Туг | Lys 205 | Thr | Gly | Ser |
| 10 | Asn | Asn 210 | Phe | Leu | Ser | Leu | S∈r 215 | Lys | Leu | Glu | Glu | Ala 220 | Ile | Val | Asp | Val |
| | Asn 225 | Ala | Thr | Lys | Aεn | Lys 230 | LYS | Het | Γ^{j} . ϵ | Glγ | Ser 235 | Asp | Tyr | Thr | Val | Ala 240 |
| | Thr | Glу | Arg | Leu | Arg 245 | Phe | G1 y | Ile | Asp | Phe 250 | Thr | Pro | Glu | Trp | Ser 255 | Leu |
| 15 | | | | 260 | | | | | 265 | | | | | 270 | Glγ | |
| | Ser | lle | Trp 275 | G17. | Val | Туr | G17. | Glu 280 | Ser | Γ/.ε | L'.'E | Asn | Leu 285 | Asn | Arg | Ser |
| 20 | Ser | Thr 290 | Ser | Phe | Glu | Leu | Leu 295 | G17. | Γ.'.ε | His | G1; | Суs 300 | His | Thr | Leu | Gln |
| | Phe 305 | Ser | Pro | Tyr | Phe | Asn 310 | Ile | Glu | Lys | Ser | Glu 315 | Asn | Tyr | Asn | Asn | Ala 320 |
| | Asp | Pro | Thr | Glγ | Phe 325 | Ile | Asn | Туг | Γλε | Ser 330 | Asp | T;r | Tyr | Thr | Tyr 335 | Glу |
| 25 | | | | 340 | | | | | 345 | | | | | 350 | Val | |
| | GJ. | Val | Դ≈բ 355 | Ser | Arg | Asn | Het | Thr 360 | Het | Glu | Ser | Glu | Arg 365 | Phe | Glu | Gln |
| 30 | Ala | G17 370 | Val | Asn | Thr | Lys | Pro 375 | Тут | Asn | Pro | G17 | T;r 380 | Ala | Thr | Asn | Asn |
| | 11≘ 385 | Gl7 | Leu | Phe | Gly | Gln 390 | Ala | Asn | Phe | Tyr | Leu 395 | L⊕u | Asn | Asp | Ala | Leu 400 |
| | Ser | Ile | Ser | Ala | Gly 405 | Ala | Arg | Ala | Asp | Phe 410 | Het | Phe | Phe | Asp | Leu 415 | Lys |
| 35 | Alā | Asn | Glu | Tyr 420 | Leu | Asn | Åsn | Glu | Ala 425 | Lys | Gln | Glu | Thr | His 430 | Asn | Val |
| | Ile | Asn | Pro 435 | Asn | Val | Gly | Ile | Lys 440 | Туг | Glu | Phe | Val | Lys 445 | Gly | Leu | Thr |
| 40 | Ala | His 450 | Gl y | Thr | Phe | Gly | Ser 455 | Ala | Phe | Ser | Ala | Pro 460 | Asp | Ala | Ph⊜ | Gln |
| | 465 | | | | | 470 | | | | | 475 | | | | Asn | 480 |
| 45 | A£p | Leu | Γλε | Pro | Glu 485 | Lys | Ser | Met | Thr | Trp 490 | Asp | Phe | Gly | Ile | Gly 495 | Туг |
| 45 | | | | 500 | - | | | | 505 | • | | | | 510 | Туr | |
| | | | 515 | | | | | 520 | | | | | 525 | | Ala | |
| 50 | | 530 | | | | | 535 | | | | | 540 | | | G17. | |
| | 545 | | | | | 550 | | | - | | 555 | | | | Lys | 560 |
| 5 5 | | | | | 565 | | | | | 570 | | | | | Glu 575 | |
| 33 | | | | 580 | | | | | 585 | | | | | 590 | Val | |
| | | | 595 | | | | | 600 | | | | | 605 | | | Leu |
| 60 | | 610 | | | | | 615 | | | _ | | 620 | | | Gln | |
| | 625 | | | | | 630 | | | | | 635 | | | | Leu | 640 |
| e e | | | | | 645 | | | | | 650 | | | | _ | His 655 | |
| 65 | | | | 660 | | | | | 665 | | _ | | | 670 | ГЛ2 | _ |
| | | | 675 | | | | | 680 | | | | - | 685 | | Tyr | |
| 70 | | 690 | | | Tyr | His | 11et 695 | Pro | Gl; | Arg | Asn | 11e 700 | Het | Glγ | Lys | Val |
| | 11et 705 | Vāl | Asn | Phe | | | | | | | | | | | | |
| <i>7</i> 5 | (2) | INFO | ORI IAC | [1011 | FOR | SEQ | IDI | 10:4 | 19 | | | | | | | |

```
(i) SEQUENCE CHARACTERISTICS:
                   (A) LEHGTH: 462 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
 5
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
10
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
15
                   (B) LOCATION 1...462
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:449
       Net Asn Arg Phe Ser Asn His Trp Pro Cys Ile Leu Val Gly Phe Val
20
       Leu Trp Phe Val Ser Ala Ser Arg Thr Val Ala Gln Asn Ala Ser Glu
20 25 30
       Thr Thr Val Ser Tyr Asp Thr Asp Thr Ala Val Leu Ser Glu Ala Asp
35 40
       Val Leu Arg Ile Ala Leu Ser Glu Asn Ala Thr Val Lys Val Ala Asp
50 60
       Het Asp Val Arg Lys Gln Glu Tyr Ala Arg Arg Ala Arg Ala Asp
65 70 75 80
       Leu Phe Pro Lys Val Asp Leu Asn Gly Val Tyr Ser His Thr Leu Lys
85 90 95
30
       Lys Gln Val Leu Tyr Ile Asp Het Pro Gly Phe Ser Ser Glu Gly 100 105 110
       Ile Glu Met Gly Arg Thr His Asn Thr Gln Gly Gly Val Asn Val Ser 115 120 125
       Het Pro Leu Val Ser Ala Gln Leu Trp Lys Ser Ile Ala Met Thr Gly
130 135 140
35
       Glu Gln Leu Asp Leu Ala Leu Glu Lys Ala Arg Ser Ser Arg Ile Asp
145 150 155
       Leu Val Ala Glu Val Lys Lys Ala Tyr Leu Ser Val Leu Leu Ala Glu
165 170
40
       Asp Ser Tyr Gly Val Phe Lys Arg Ser Tyr Asp Asn Ala Leu Ala Asn
180 185 190
       Tyr Lys Asn Ile Ser Asp Lys Phe Asp Arg Gly Leu Val Ala Glu Tyr
195 200 205

    Asp
    Lys
    Ile Arg
    Ala
    Asn
    Val
    Gln
    Val
    Arg
    Asn
    Ile Glu
    Pro
    Asn
    Leu

    Leu
    Gln
    Ala
    Gln
    Asn
    Ser
    Val
    Ala
    Leu
    Ala
    Leu
    Trp
    Gln
    Leu
    Lys
    Val

    225
    230
    235
    240

45
       Leu Het Ser Met Glu Val Glu Thr Pro Ile Arg Leu Ser Gly Ser Leu
245 250 255
50
       Ser Asp Tyr Lys Glu Gln Val Tyr Thr Gly Tyr Phe Ala Ala Asp Thr
260 265 270
       Leu Ile Ser Asn Asn Ser Ser Leu Arg Gln Leu Asp Ile Gln Arg Arg
275 280 285
       Leu Ala Val Ser Ala Asp Lys Leu Asn Lys Tyr Ser Phe Leu Pro Thr
290 295 300
55
        Leu Asn Leu Gly Gly Gln Tyr Thr Tyr Ser Leu Asn Ser Asn Asp Ile
305 310 315 320
        Lys Phe Trp Gly Glu Gly Gln Arg Trp Thr Pro Phe Ser Thr Ile Ser
325 330 335
60
        Leu Ser Leu Tyr Ile Pro Ile Phe Asn Gly Gly Lys Arg Leu Tyr Asn 340 345 350
        Val Lys Gln Ser Ala Leu Ser Ile Arg Gln Ile Asp Leu Gln Arg Arg
355 360 365
        His Ile Glu Gln Ser Ile Arg Met Gly Ile Lys Asn Gln Asn Asp Arg
370 375 389
65
        Leu Arg Thr Cys Het Gln Arg Phe Val Ala Ser Glu Glu Ala Val Arg
385 390 395
        Ser Ala Glu Lys Gly Tyr Gln Ile Ala Glu Lys Arg Tyr Gln Thr Gly 405 415
70
        Glu Gly Thr Leu Val Glu Leu Asn Asp Ala Asp Val Ala Leu Leu Gln
420 430
        Ala Arg Leu Asn Tyr Asn Gln Ala Ile Phe Asp Phe Het Thr Ala Lys 435 440 445
                                       440
        Ala Glu Leu Asp Lys Met Asn Gly Met Gly Ile Pro Glu Gln
75
```

```
450
                                   455
                                                          460
       (2) INFORMATION FOR SEQ ID NO:450
 5
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 492 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
10
            (ii) HOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
15
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...492
20
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:450
       Het Trp Gly Asp Ser His Gly Val Ala Pro Asn Gln Val Arg Arg Thr
                                                10
       Leu Val Lys Val Ala Leu Ser Glu Ser Leu Pro Fro Gly Ala Lys Gln
20 25 30
       Ile Arg Ile Gly Phe Ser Leu Pro Lys Glu Thr Glu Glu Lys Val Thr 35 40 45
       Ala Leu Tyr Leu Leu Val Ser Asp Ser Leu Ala Val Arg Asp Leu Pro 50 60
30
       Asp Tyr Lys Gly Arg Val Ser Tyr Asp Ser Phe Pro Ile Ser Lys Glu 65 70 75 80
       Asp Arg Thr Thr Ala Leu Ser Ala Asp Ser Val Ala Gly Arg Arg Phe 85 90 95
35
       Phe Tyr Leu Ala Ala Asp Ile Gly Pro Val Ala Ser Phe Ser Arg Ser
100 105 110
       Asp Thr Leu Thr Ala Arg Val Glu Glu Val Ala Val Asp Gly Arg Pro
115 120 125
       Leu Pro Leu Lys Glu Leu Ser Pro Ala Ser Arg Arg Leu Tyr Arg Gly
130 135 140
40
       Tyr Glu Ala Leu Phe Val Pro Gly Asp Gly Gly Ser Arg Asn Tyr Arg
145 150 155 160
       Ile Pro Ala Ile Leu Lys Thr Ala Asn Gly Thr Leu Ile Ala Net Ala
165 170 175
       Asp Arg Arg Lys Tyr Asn Gln Thr Asp Leu Pro Glu Asp Ile Asp Ile
180 190
45
       Val Het Arg Arg Ser Thr Asp Gly Gly Lys Ser Trp Ser Asp Pro Arg
195 200 205
       Ile Ile Val Gln Gly Glu Gly Arg Asn His Gly Phe Gly Asp Val Ala
210 215 220
50
       Leu Val Gln Thr Gln Ala Gly Lys Leu Leu Met Ile Phe Val Gly Gly 225 230 235 240
       Val Gly Leu Trp Gln Ser Thr Pro Asp Arg Pro Gln Arg Thr Tyr Ile 245 250 255
       Ser Glu Ser Arg Asp Glu Gly Leu Thr Trp Ser Pro Pro Arg Asp Ile
260 265 270
55
       Thr His Phe Ile Phe Gly Lys Asp Cys Ala Asp Pro Gly Arg Ser Arg 275 280 285
       Trp Leu Ala Ser Phe Cys Ala Ser Gly Gln Gly Leu Val Leu Pro Ser 290 295 300
60
       Gly Arg Ile Thr Phe Val Ala Ala Ile Arg Glu Ser Gly Gln Glu Tyr
305 310 315 320
       Val Leu Asn Asn Tyr Val Leu Tyr Ser Asp Asp Glu Gly Asp Thr Trp
325 330 335
65
       Gln Leu Ser Asp Cys Ala Tyr Arg Arg Gl_{Y} Asp Glu Ala Lys Leu Ser 340 345 350
       Leu Met Pro Asp Gly Arg Val Leu Met Ser Ile Arg Asn Gln Gly Arg 355 360 365
       Gln Glu Ser Arg Gln Arg Phe Fhe Ala Leu Ser Ser Asp Asp Gly Leu
370 375 380
70
       Thr Trp Glu Arg Ala Lys Gln Phe Glu Gly Ile His Asp Pro Gly Cys
385 390 395 400
       Asn Gly Ala Het Leu Gln Val Lys Arg Asn Gly Arg Asp Gln Val Leu
405 410 415
75
       His Ser Leu Pro Leu Gly Pro Asp Gly Arg Arg Asp Gly Ala Val Tyr
```

```
420
                                            425
       Leu Phe Asp His Val Ser Gly Arg Trp Ser Ala Pro Val Val Val Asn 435 440 445

Ser Gly Ser Ser Ala Tyr Ser Asp Met Thr Leu Leu Ala Asp Gly Thr 450 455 460
       Ile Gly Tyr Phe Val Glu Glu Gly Asp Glu Ile Ser Leu Val Phe Ile 465 475 480
       470 475
Arg Phe Val Leu Asp Asp Leu Phe Asp Val Arg Gln
10
       (2) INFORMATION FOR SEQ ID NO: 451
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 245 amino acids
15
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
20
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
25
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...245
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:451
30
       Met Lys Lys Glu Lys Leu Trp Ile Ala Ile Val Ala Gly Leu Ala Phe
                                                10
       Val Leu Gly Leu Tyr Ala Leu Gly Arg Ser Val Ala Gln Leu Arg Arg 20 25 30
35
       Ser Gln Pro Ser Val Thr Val Thr Gly Het Ala Glu Arg Asn Phe Lys
       Ser Asp Leu Ile Val Trp Thr Ala Ser Tyr Gln Leu Gln Het Het Asp 50 60
       Leu Glu Ser Ala Tyr Lys Ala Leu Lys Glu Lys Gln Ile Leu Val Ala
65 70 75 80
40
       Asp Tyr Leu Lys Asn Lys Gln Leu Pro Asp Ser Ser Tyr Ile Phe Ser
85 90 95
       Ser Val Ala Ile Ser Lys Glu Tyr Asn Tyr Tyr Tyr Asp Pro Arg Gln 100 105 110
       Glu Gln Asn Val Arg Thr Phe Ala Gly Tyr Leu Leu Ser Gln Thr Val
115 120 125
45
       Thr Val Thr Ser Gln Asp Ile Glu His Val Glu Lys Ile Ser Arg Asp
130 135 140
       Ile Thr Glu Leu Ile Asn Gln Gly Val Glu Ile Thr Ser Asp Arg Pro
145 150 155 160
50
       Ala Tyr Tyr Tyr Thr Lys Leu Asn Asp Leu Lys Val Glu Met Leu Arg
165 170 175
       Asn Ala Ser Glu Asp Ala Phe Asn Arg Ala Ser Val Ile Ala Glu Gly
180 185 190
55
       Ser Gly Ser Ser Val Gly Lys Met Leu Ser Ser Ser Met Gly Val Phe
                                     200
       Gln Ile Val Gly Leu Asn Ser Asn Glu Asp Tyr Ser Trp Gly Gly Ser
210 215 220
       Phe Asn Thr Ser Ser Lys Het Lys Thr Ala Ser Ile Thr Val Lys Ala
60
       Ser Phe Ala Leu Lys
       (2) IMPORNATION FOR SEQ ID NO: 452
65
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 276 amino acids
                  (B) TYPE: amino acid
(D) TOPOLOGY: linear
70
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
75
            (vi) ORIGINAL SOURCE:
```

WO 99/29870

```
(A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
 5
                   (B) LOCATION 1...276
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:452
       Het Lys Lys Thr Ile Ala Ile Ile Ala Ser Ala Leu Leu Ala Leu Gly 1 5 10 15
10
       Ala Val Gl; Cys Lys Lys Asn Ala Asp Thr Thr Ala Val Ser Glu Lys
20 25 30
       Asp Ser Ile Ala Leu Ser Met Gly Ile Leu Tyr Gly Gln Asp Phe Ala 35 40 45
       Asn Gln Phe Glu Net Ser Arg Leu Gln Gly Gln Pro Ile Asp Ser Val
50 55 60
15
       Ala Phe Leu Asp Gly Phe Lys Tyr Gly Ile Asp Thr Thr Arg Phe Ser 65 70 75 80
       Tyr Asn Leu Gly Ala Ile Tyr Ala Ser Asn Ile Ala Arg Gln Leu Ala
85 90 95
20
       His Asp Ser Ile Asp Ile Asp Lys Fhe Tyr Ala Ala Het Arg Ala Ala 100 105 110
       Leu Leu Lys Asp Thr Val Ser Ile Ala Het Lys Pro Ala Asp Ala Gln 115 120 125
25
       Ala Phe Het Gln Arg Ile Gln Ala Lys Lys Gln Arg Glu Asn Asn Met
130 135 140
       Lys Gln Phe Gly Gln Asn Ile Glu Lys Gly Asn Glu Tyr Ile Asp Thr
145 150 160
       Phe Lys Lys Glu Asp Gly Val Thr Val Thr Thr Gly Leu Ala Tyr 165 170 175
30
       Lys Thr Leu Gln Glu Gly Thr Gly Ala Thr Pro Ser Leu Ala Asp Thr 180 185
       Val Arg Val Lys Tyr Val Gly Thr Leu Val Asp Gly Lys Glu Phe Asp
195 200 205
35
       Lys Asn Glu Glu Gly Ile Glu Phe Ala Val Thr Gly Val Ile Lys Gly 210 215 220
       Trp Thr Glu Net Leu Gln Leu Met Lys Val Gl; Gln Lys Val Arg Val 225 230 235 240
       Val Ile Pro Gln Glu Leu Ala Tyr Gly Glu Thr Gly Asn Tyr Thr Ile
245 250 255
40
       Glu Pro Phe Ser Thr Leu Thr Phe Glu Het Glu Leu Ile Gly Ile Lys
                   260
       Pro Gly Lys Lys
45
       (2) INFORMATION FOR SEQ ID NO:453
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 775 amino acids
50
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
55
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
60
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...775
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 453
65
       Het Lys Val Leu Arg Gln Val Phe Leu Pro Ile Leu Phe Val Leu Leu 1 \phantom{\bigg|}5\phantom{\bigg|} 10 \phantom{\bigg|}15\phantom{\bigg|}
       Thr Gly Ala Cys Ser Thr Thr Lys Asn Leu Pro Glu Gly Glu Gln Leu 20 25 30
70
       Tyr Ile Gly Het Gly Lys Thr Gln Ile Leu Arg Gln Asp Lys Ser His 35 40 45
                                                             4.5
       Ala Gly Gln Gln Ala Leu Thr Glu Val Glu Ser Thr Leu Lys Val Thr
                                55
       Pro Ash Gly Ala Ile Phe Gly Ser Ala Ser Ala Ser Leu Pro Lys Ile
75
```

| | Pro | Ph∈ | GJ ?. | Leu | Trp | Leu | Tyr | Asn | Ser | Phe | Val | G1 ?. | Asp | Ser | Thr | Vāl |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------------|------------|
| | lle | Ser | L∵s | Trp | 85 Ile | Fhe | Asp | Lys | Phe | 90 Ala | Ala | Lys | Pro | Val | 95 Phe | Ile |
| 5 | | | | 100 | | | | Arg | 105 | | | | | 110 | | |
| Ū | | | 115 | | | | | 120 Ala | | | | | 125 | | | |
| | | 130 | | | | | 135 | | | | | 140 | | | | |
| 10 | 145 | | | | | 150 | | Ala | | | 155 | | | | | 160 |
| | | | | | 165 | | | Ser | | 170 | | | | | 175 | |
| | Phe | Pro | λερ | Ser 180 | Ile | Leu | Ala | Tyr | Arg 185 | Gln | Thr | Pro | Ser | Leu 190 | Ile | Arg |
| 15 | Γλε | Glу | Asp 195 | Gln | Phe | Asn | Leu | Ala 200 | | Leu | His | Glu | Glu 205 | Arg | Gln | Thr |
| | Ile | | Ala | Leu | Leu | Arg | Asp | Asn | Gly | Tyr | Tyr | T;r 220 | | Arg | Pro | Gln |
| 00 | | 210 Ile | Ile | T;:r | Glu | | 215 Asp | Thr | Leu | Leu | Val | | Gly | Ala | Val | C;;s |
| 20 | 225 Leu | Arg | Alā | Lys | Leu | 230 Ser | Glu | Aεp | Thr | | 235 Pro | Gln | Ala | Het | Arg | 240 Pro |
| | Trp | Arg | Ile | Gly | 245 Lys | Arg | Thr | Ala | Val | 250 Leu | Leu | Gl y | Met | Asn | 255 Gly | Glu |
| 25 | | | | 260 | | | | Val | 265 | | | | | 279 | | |
| | | | 275 | | | | | 280 Lys | | | | | 285 | | | |
| | | 290 | | | | | 295 | Gln | | | | 300 | | | | |
| 30 | 305 | | | | | 310 | | | | | 315 | | | | | 320 |
| | | | | | 325 | | | Phe | | 330 | | | | | 335 | |
| | | | | 340 | | | | Leu | 345 | | | | | 350 | | |
| 35 | | | 355 | | | | | Ser 360 | | | | | 365 | | | |
| | Ser | Asn 370 | | Fhe | Ile | Gl y | Pro 375 | Gly | Leu | Asn | Phe | Ala 380 | Leu | Ala | Arg | Arg |
| 40 | Asn 385 | Val | Phe | Gly | Gly | Gly 390 | Glu | Asn | Leu | Ser | Trp 395 | Asn | Ile | Gly | Gly | Ser 400 |
| 10 | Tyr | Glu | Trp | Glu | | | | Arg | Pro | Glu 410 | Asn | Ser | Ser | Asn | Arg 415 | Leu |
| | Ile | Asp | Ile | | | Tyr | Asn | Met | Asn | Thr. | | Val | Asn | Leu 430 | | Phe |
| 4 5 | Pro | Ser | | 420 Val | | Pro | Gly | Leu | 425 Leu | | Lys | Tyr | | Tyr | Tyr | Pro |
| | Thr | Thr | 435 Thr | Thr | Phe | Gln | | 440 Ser | Alā | Thr | Ala | | | | Ala | His |
| | Tyr | 450 Phe | | Het | Tyr | Ser | 455 Phe | Gly | Phe | Ser | Thr | 460 Thr | | Glu | Phe | Gln |
| 50 | 465 | | | | | 470 | | Ala | | | 475 | | | | | 480 |
| | | | | | 485 | | | Glu | | 490 | | | | | 495 | |
| 55 | | | | 500 | | | | Leu | 505 | | | | | 510 | | |
| 33 | | | 515 | | | | | 520 | | | | | 525 | | | |
| | | 530 | | | | | 535 | | | | | 540 | | | | |
| 60 | 545 | | | | | 550 | | Leu | | | 555 | | | | | 560 |
| | Leu | Ile | Tyr | Leu | Ala 565 | | Gl y | Lys | Lys | Tyr 570 | | Asp | Thr | Lγε | A <i>≘</i> n 575 | Phe |
| | Vāl | Gly | Val | Pro 580 | | Ser | Gln | Phe | 11e 585 | | Ala | Thr | Gly | Glu 590 | Leu | Arģ |
| 65 | туг | Ser | Tyr 595 | Thr | | Asp | Arg | Asn 600 | | Ser | Leu | Ala | Thr | | Phe | Gly |
| | Thr | | Val | | Tyr | Ser | | Gly | | Het | Arg | Val 620 | Ala | | T;r | Ser |
| 70 | | | | туг | Val | | | Ala | Asn | Ser | | Arg | | Phe | Thr | Val |
| 70 | 625 Arg | , J Ser | Ile | Gly | | | | , Fhe | Asn | | | | Asp | Asn | | |
| | Ser | туг | Leu | Агр | 645 Gln | | Gl | . Glu | Phe | 650 Lys | | Glu | Ala | . Asn | 655 Val | Glu |
| 75 | | - | | 660 |) | | | | 665 | , | | | | 670 | | Asp |
| - | | | | -,- | | | . – , | | | | | | | | | • |

```
680
                                                               685
       Ala Gly Asn Val Trp Leu Leu Arg Glu Asp Ser Ser Arg Pro Gly Gly 690 695 700

Ala Leu Ser Glu Val Gly Ser Val Ser Asn Phe Leu Asn Ser Ile Ala 715 720
 5
       Leu Gly Thr Gly Val Gly Leu Arg Tyr Asp Leu Ala Phe Leu Val Val 725 730 735
       Arg Val Asp Val Gly Phe Gly Leu His Leu Pro Tyr Asn Thr Gly Lys 740 745 750
10
       Lys Gly Tyr Tyr Asn Ile Pro Arg Phe Lys Asp Ala Ile Gly Phe His
               755
       Leu Ala Val Gly Tyr Pro Phe
15
       (2) INFORMATION FOR SEQ ID NO: 454
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 774 amino acids
                   (B) TYPE: amino acid
20
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
25
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
30
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...774
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:454
35
       Leu Leu Het Leu Ala Ser Ser Cys Ser Val Thr Arg Tyr Val Pro Asp
20 25 550
       Gly Ser Arg Leu Leu Asp Arg Val Thr Ile Ala Ser Glu Thr Gly Ser 35 40 45
40
       Ile Ala Leu Pro Glu Asp Ile Arg Asp Tyr Thr Leu Gln Gln Pro Asn 50 55 60
       Tyr Arg Leu Phe Gly Met Thr Arg Trp Leu Leu Arg Val Tyr Ser Ser 65 70 75 80
45
       Ser Asn Pro Asn Ser Asn Ser Trp Trp Asn Arg Ser Leu Arg Lys Met 85 90 95
       Gly Glu Pro Pro Val Leu Ile Asp Ser Val Leu Thr Asp Arg Thr Ala 100 105 110
       Asn Arg Leu Ala Lys Ala Met Ala Gly Asp Gly Phe Leu Asp Ala Thr
115 120 125
50
       Ala Arg Ala Val Val Asp Thr Gly Leu Tyr Lys Lys Ala Arg Ile Thr
130 135 140
       Tyr Leu Ile Gln Fro Gly Ser Arg Tyr Tyr Ile Arg Asn Net Ala Leu 150 150 155 160
55
       Asp Val Lys Asn Pro Leu Leu Pro Pro Val Ala Leu Gly Asn Ser Leu
165 170 175
       Pro Ser Ala Tyr Lys Val Gly Ile Ser Glu Gly Ser Pro Leu Ser Pro
180 185 190
       Ile Val Leu Asp Glu Glu Arg Lys Ala Ile Ala Arg His Net Arg Asn 195 200 205
60
       Asn Gly Phe Trp Lys Phe Ser Ala Glu Asp Val Tyr Glu Ala Asp 210

Thr Thr Val Ser Gly Gly Ser Gly Thr Lys Ser Ala Asp Leu Lys Leu 225

230

240
65
       Val Val Asn Gly Ile Gly Arg Tyr Pro Tyr Arg Ile Gly Arg Val Phe
245 250 255
       Phe His Ala Asp Tyr Asp Pro Leu Glu Ser Asp Phe Arg Val Gln Glu 260 265 270
       Leu Pro Arg Ile Asp Ser Ile Ser Arg Gly Asp Tyr Thr Val Tyr Tyr 275 280 285
70
       Gly Ser Arg Gly Arg Tyr Ile Arg Ala Ser Ala Leu Thr Arg Ser Val
290 295 300
       Ser Val Thr Pro Gly Ala Phe Phe Cys Glu Asp Asp Val Glu Arg Ser
                              310
                                                   315
       Tyr Ile Lys beu Asn Ala Leu Pro Ile Val Arg Asn Val Asn Ile Arg
```

| | 325 330 335 Phe Val Glu His Asn Gly Lys Asp Glu Ile Ala Leu Ala Asp Ser a | | | | | | | | | | | | | | Ser | |
|------------|--|------------|------------|------------|----------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|
| | | Leu | | 340 | | | | | 345 | | | | | 350 | | |
| 5 | | Glu | 355 | | | | | 360 | | | | | 365 | | | |
| | | 370 Leu | | | | Phe | 375 | | | | Leu | 380 | | | | Glu |
| 10 | 385 Het | Phe | Asn | Ile | | 390 Leu | Lys | Glу | Ala | Tyr | 395 Glu | Ala | Ile | Arg | Lys 415 | 400 Gl ; |
| | Ser | His | Ser | Phe 420 | 405 Met | Glu | Tyr | Glγ | Val 425 | 410 Glu | ser | Ser | Leu | Arg 430 | | Pro |
| 15 | Arg | Leu | Leu 435 | Phe | Pro | Phe | Ile | Ser 440 | Asp | Glu | Thr | Arg | Arg 445 | Arg | Leu | Arg |
| 10 | | Ser 450 | Thr | | | | 455 | Glү | | | | 460 | | | | |
| | 465 | Asp | | | | 470 | | | | | 475 | | | | | 480 |
| 20 | | Leu | | | 485 | | | | | 490 | | | | | 495 | |
| | | Leu | | 500 | | | | | 505 | | | | | 510 | | |
| 25 | | Thr Tyr | 515 | | | | | 520 | | | | | 525 | | | |
| | | 530 Asn | | | | | 535 | | | | | 549 | | | | |
| 30 | 545 | Gln | | | | 550 | | | | | 555 | | | | | 560 |
| | | Туr | | | 565 | | | | туг | 570 | | | | Lys | 575 | |
| 0.5 | | Asp | Leu | 580 | | | | Leu | 585 | | | | Asn | 590 | | |
| 3 5 | Leu | His | 595 Leu | Gly | Phe | Gly | Leu 615 | 600 Ala | Phe | Pro | Туr | Gly 620 | 605 Asn | Ala | Arg | His |
| | Ile 625 | 610 Pro | Phe | Glu | Leu | Arg 630 | Tyr | Phe | Ala | Glγ | G1; 635 | | Asn | Ser | Vāl | Arg 640 |
| 4 0 | G1 y | Trp | Ser | Val | Arg 645 | Thr | Leu | Gly | Pro | G1; 650 | Ser | Het | Lys | Met | Thr 655 | Pro |
| | | Γ'ε | | 660 | Phe | | | | 665 | | | | | 670 | | |
| 4 5 | | Glu | 675 | | | | | 680 | | | | | 685 | | | |
| | | Asp 690 | | | | | 695 | | | | | 700 | | | | |
| 50 | 705 | G1; | | | | 710 | | | | | 715 | | | | | 720 |
| 30 | | Ala | | | 725 | | | | | 730 | | | | | 735 | |
| | | Ala | | 740 | | | | | 745 | | | | | 750 | | |
| 55 | | Val | 755 | | | | | 760 | | | | | 765 | | | |
| | | 770 | | | | | | | | | | | | | | |
| 60 | (2) | 111F | | | | | | | | | | | | | | |
| | | (1 | | A) L | CE C ENGT TPE: | H: 8 | 67 a | mino | | ds | | | | | | |
| 65 | | | | | OPOL | | | | | | | | | | | |
| | | (ii |) 110 | LECU | LE T | YPE: | pro | tein | | | | | | | | |
| | | • | .) HY | | | | | | | | | | | | | |
| 70 | (∵i) ORIGINAL SOURCE: (A) ORGANISH: Porphyromonas gingivalis | | | | | | | | | | | | | | | |
| | | (is | :) FE | | | WEV. | mia | c f | s.+ | ٠. | | | | | | |
| 75 | | | (| B) I | AME/ OCAT | NOI. | 1 | 867 | acul | . = | | | | | | |

| (xi) SEQU | ENCE | DESCRIPTION: | SEQ | ΙD | NO:455 |
|-----------|------|--------------|-----|----|--------|
|-----------|------|--------------|-----|----|--------|

| 5 | Net 1 | Arg | Lys | Arg | Ile 5 | Leu | Gln | Leu | Phe | Leu 10 | Thr | Ala | Leu | Leu | Leu 15 | Ala |
|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------------------|------------|------------|
| | | Gly | Ser | Ser 20 | Leu | Ala | Ile | Ala | Gln 25 | | Val | Val | Thr | Gly 30 | Lys | Val |
| | Ile | Asp | Ser 35 | Glu | Thr | Ser | Glu | Pro 40 | Leu | Ile | Gly | Vāl | Ser 45 | Val | Ser | Thr |
| 10 | G1; | Gln 50 | G17 | Ala | Ser | Leu | Arg 55 | Gl? | Val | Thr | Thr | Asp 60 | Het | Asp | Gly | G17. |
| | Phe 65 | Arg | Phe | Glu | Val | Pro 70 | Ala | Γλε | Ser | Val | Leu 75 | Thr | Phe | Arg | Cle | Val |
| 15 | G1 ?. | Tyr | Ala | Thr | Val 85 | Thr | Arg | Ser | Ile | Gly 90 | Arg | Glγ | Ser | Gln | Glu 95 | Asp |
| | Leu | G1;. | Thr | 11e 100 | Leu | Leu | Asp | Pro | Gln 105 | Ala | Ile | G1y | Leu | Asp 110 | Glu | Ile |
| | Gln | Vál | Ile 115 | Ala | Ser | Vāl | Val | Pro 120 | Lys | Asp | Arg | 11e t | Thr 125 | Pro | Val | Pro |
| 20 | Vāl | Ser 130 | Asn | Ile | Arg | Val | Ala 135 | Asp | Ile | Gln | Alā | Ala 140 | Ser | Leu | Asn | Val |
| | Glu 145 | Phe | Pro | Glu | Leu | Val 150 | Lys | Ser | Thr | Pro | Ser 155 | Thr | Туг | Thr | Thr | Lys 160 |
| 25 | Glγ | Ser | G17. | G17. | Phe 165 | Gl; | Asp | Gly | Arg | Thr 170 | Asn | Val | Arg | Glγ | Phe 175 | Asp |
| | Thr | Tyr | Asn | Phe 180 | Gl y | Val | Leu | Ile | Asn 185 | Gl; | Val | Pro | Val | Asn 190 | G17. | Het |
| 00 | | | 195 | - | | | - | 200 | | - | | - | 205 | | Asn | |
| 30 | Alā | Ser 210 | Thr | Ile | Gln | Ile | Gl n 215 | Arg | Glγ | Leu | G17 | Ala 220 | Ser | Lys | Leu | Gl 7 |
| | 11e 225 | Ser | Ser | Väl | Glγ | Gly 230 | Thr | Met | Asn | Ile | 11e 235 | Thr | Lys | Thr | Thr | Asp 240 |
| 35 | | | | | 245 | | | | | 250 | | _ | | _ | Gly 255 | |
| | His | Lγε | Glu | Ser 260 | Phe | Ser | Ile | Ser | Thr 265 | Gl; | Het | Asn | Ағр | Gl _Y 270 | Trp | Ala |
| 40 | | | 275 | | | | | 280 | | | | | 285 | | ГЛЕ | |
| 40 | | 290 | | | | | 295 | | | | | 300 | | | Γλε | |
| | Asn 305 | Glu | Arg | His | Thr | Leu 310 | Ser | Leu | Thr | Gly | Phe | G17. | Ala | Pro | Gln | Trp 320 |
| 45 | | | | | 325 | | | | | 330 | | | _ | | Lys 335 | _ |
| | | | | 340 | | | | | 345 | _ | | | _ | 350 | Leu | |
| 50 | | | 355 | | | | | 360 | | | | | 365 | | Ser | |
| 50 | | 370 | | | | | 375 | | | | | 380 | - | | Ala | |
| | 385 | | | | | 390 | | | | | 395 | | | | Lys | 400 |
| 55 | | | | | 405 | | | | | 410 | | | | | Glu 415 | |
| | | | | 420 | | | | | 425 | | | | | 430 | Leu | |
| 60 | | | 435 | | | | | 440 | | | | | 445 | | Leu | |
| 00 | | 450 | | | | | 455 | | | | | 460 | | | Lγε - | |
| | 465 | | | | | 470 | | | | | 475 | | | | T/r | 480 |
| 65 | | | | | 485 | | | | | 490 | | | | | Ser 495 | |
| | | | | 500 | | | | | 505 | | | | | 510 | Gly | |
| 70 | | | 515 | | | | | 520 | | | - | _ | 525 | | Glu | |
| , 0 | | 530 | | | | | 535 | | | | | 540 | | | Trp | |
| | 545 | | | | | 550 | | | | | 555 | | | | Asn | 560 |
| 7 5 | Asn | Tyr | GI; | Gly | Ser 565 | Γ'.ε | Ser | Thr | Glу | Tyr 570 | Leu | Pro | Glу | Val | Ser 575 | Pro |
| | | | | | | | | | | | | | | | | |

```
Trp Lys Ser Phe Leu Pro Trp Ser Gly Lys Ala Gly Leu Ser Tyr Lys
580 590
       Phe Ala Gln Gly His Asn Val Phe Ala Asn Gly Gly Fhe Phe Thr Arg
595 600 605
      Ala Pro Leu Phe Gly Asn Ile Tyr Ala Ala Gl; Ala Ile Ile Pro Asn
610 615 620
       Asp Lys Ala Ash Het Glu Lys Val Leu Thr Gly Glu Val Gly Tyr Gly
625 630 635 640
       Phe Thr Ash His Lys Ash Phe Glu Phe Ash Ile Ash Gly Tyr Tyr Thr
645 650 655
10
       Lys Trp Het Asp Arg Val Thr Ser Lys Arg Ile Gly Asn Glu Tyr Val
660 665 670
       Tyr Leu Asn Giy Val Asp Ala Val His Cys Gly Val Glu Ala Glu Val
675
680
685
Ser Tyr Arg Pro Ile Arg Gln Ile Asp Leu Arg Gly Het Phe Ser Leu
690
695
15
       Gly Asp Trp Thr Trp Gln Asn Asn Val Ser Tyr Thr Ser Tyr Asp Glu
705 710 715 720
       Ala Gly Asn Glu Thr Gly Gln Asp Ile Thr Tyr Ile Lys Gly Leu His 725 730 735
20
       Val Gly Asp Ala Ala Gln Het Thr Ala Ala Val Ser Ala Asp Ile Glu
740 750
       Leu Phe Lys Gl; Phe His Val Ile Gly Lys Tyr Asn Phe Leu Gly Lys
765 765
       Ash Tyr Ala Gly Phe Ash Pro Ala Thr Arg Ash Ala Gln Gln Tyr Glu
770 786
25
       Ala Asp Gly Lys Glu Ile Val Glu Ser Trp Lys Leu Pro Asp Val Gl; 785 790 795
       Leu Phe Asp Leu Ser Ala Ser Tyr Asn Phe Lys Leu Gly Ser Leu Ser
805 810
30
       Thr Thr Phe Tyr Phe Asn Met Asp Asn Val Ala Asp Lys Arg Tyr Val 820 825
       Ser Asp Ala Asp Asp Asn Ile Ile Gly Lys Lys His Asp Glu Ala Ser
835 840 845
       Ala Lou Val Trp Tyr Gly Phe Gly Arg Thr Trp Ser Thr Gly Ile Arg
850 855 860
35
       Val Asn Phe
       865
40
        (2) INFORMATION FOR SEQ ID NO: 456
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 431 amino acids
                    (B) TYPE: amino acid
45
                    (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
50
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
55
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...431
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 456
       Het Lys Fhe Ser Ile Arg Leu Phe Leu Cys Ile Ile Phe Leu Leu Ser 1 \phantom{-}5\phantom{+} 15
60
        Ala Phe Ile Leu Pro Ala Leu Gly Gln Lys Ser Lys Gln Val Gln Arg
20 25 30
        Leu Glu Lys Gln Arg Lys Glu Ala Leu Lys Ala Ile Glu Lys Thr Asp
35 40 45
65
        Arg Glu Leu Arg Asn Thr Lys Lys Asp Lys Gln Asp Lys Gln Lys His
50 55 60
        Leu Asn Leu Leu Asn Lys Gln Val Ala Gln Arg Lys Gln Het Val Gln
65 70 75 80
        Leu Leu Asp Asn Glu Val Lys Glu Leu Gln Ser Asp Ile Asp Ser Het
85 90
70
        Thr Gly Val Cys His Gln Leu Ser Val Glu Glu Lys Ala Arg Ser Asp
100 105 110
        Glu Tyr Ala Gln Ala Leu Gln Ser Het Gln Lys Arg Lys Arg Ser Leu
 75
```

```
Asp Arg Ile Leu Phe Ile Ser Ser Ala Lys Ser Phe Asp Glu Gly Met
130 135 140
      Arg Arg Met Arg Phe Leu Glu Gln Tyr Ala Ser Ala Tyr Lys Leu Ala
145 150 160
      Ser Val Arg Leu Arg Asp Thr Arg Ser Lys Leu Glu Thr Glu Arg Ala 165 170 175
      Thr Val Glu Asp Ala Lys Lys Glu Lys Gly His Leu Leu Val Ile Arg
180 185 190
      Glu Glu Glu Lys Lys Lys Leu Glu Gly Gln Gln Ala Glu Gln Arg
195 200 205
10
      Gln Val Gln Ala Leu Gly Ala Lys Gln Lys Asp Leu Glu Ala Gln Leu
210 215 220
      Arg Lys Gln Lys Lys Gln Ala Glu Ala Leu Asn Arg Lys Ile Glu Lys
225 230 235 240
      Gln Ile Ala Lys Glu Ile Glu Ala Ala Glu Arg Arg Ala Arg Glu Glu
245 255
15
      Arg Glu Arg Leu Ala Arg Glu Ala Lys Ala Lys Gly Lys Pro Val Pro
260 265 270
      Ala Glu Pro Glu Arg Lys Ala Glu Thr Lys Gly Gly Tyr Ala Het Asp
275 280 285
20
      Ala Ser Glu Arg Ala Leu Ser Gly Ser Phe Ala Gln Asn Lys Gly Arg
290 295 300
      25
      Ile Asp Ile Ala Val Ala Thr Gly Ser Asp Ala Thr Ser Val Phe Asp 340 345
      340 345 350
Gly Val Val Ser Ser Val Phe Val Ile Pro Gly Tyr Asn Ser Ala Val
355 360 365
30
      Het Val Arg His Gly Asn Tyr Ile Thr Val Tyr Ala Asn Leu Ser Lys
370 375 380
      Val Tyr Val Asn Ser Gly Thr Arg Val Lys Thr Gly Gln Ala Leu Gly 385 395 400
                                               395
35
      Arg Ala Tyr Thr Asp Pro Ser Asn Asn Gln Thr Ile Ile His Fhe Glu
                     405
                                        410
      Ile Trp Lys Glu Arg Ser Lys Gln Asn Pro Arg Leu Trp Leu Arg
40
      (2) INFORMATION FOR SEQ ID NO:457
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 333 amino acids
                 (B) TYPE: amino acid
45
                 (D) TOPOLOGY: linear
          (ii) NOLECULE TYPE: protein
         (iii) HYPOTHETICAL: YES
50
          (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
          (ix) FEATURE:
55
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...333
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:457
60
      Met Lys Lys Tyr Leu Leu Tyr Ala Ser Leu Leu Thr Ser Val Leu Leu
      Phe Ser Cys Ser Lys Asn Asn Pro Asn Glu Pro Val Glu Asp Arg Ser
                                    25
      Ile Glu Ile Ser Ile Arg Val Asp Asp Phe Thr Lys Thr Gly Glu Ala
65
                                40
      Val Arg Tyr Glu Arg Asn Gln Gly Ser Ala Ala Glu Arg Leu Ile Thr
                                                   60
      Asn Leu Tyr Leu Leu Phe Asp Gln Ser Gly Ala Asn Pro Ala Lys
65 70 75 80
70
      Tyr Tyr Ile Thr Gly Asn Thr Phe Thr Gly Gly Thr Trp Leu Pro Asp
85 90
      Asp Het Lys Val Lys Leu Asp Met Thr Gln Ser Glu Ala Gly Glu Arg
                 100
                                     105
      Lys Val Tyr Val Val Ala Asn Val Asp Asn Ala Val Lys Thr Ala Leu
75
```

```
Asp Ala Val Ala Asn Glu Ser Asp Leu Gln Thr Val Lys Arg Thr Thr
130 140
      Ala Net Pro Trp Ser Thr Asp Ile Ala Ser Pro Phe Leu Net Ser Gly 145 150 150
      Asn Lys Thr His Asp Phe Leu Ala Asn Arg Leu Leu Asp Asn Val Pro
165 170 175
 5
       Leu Val Arg Ala Ile Ala Lys Val Glu Leu Asn Ile Ser Leu Ser Glu
180 185 190
      Lys Phe Gln Ile Val Pro Ile Ile Val Asn Gly Ser Leu Ser Glu Phe
195 200 205
10
       Lys Phe Arg Tyr Val Asn Phe Asp Lys Glu Thr Tyr Val Val Lys Pro
       Thr Thr Lys Pro Asp Asn Leu Ile Ser Ser Ala Asn Gly Val Trp Pro
225 230 235 240
       Gln He Thr Asp Trp Thr Val Trp Gly Ala Ser Leu Asn Thr Ser Pro
245 250 255
15
       Ala Pro Asp Ala Gly Thr Gly Tyr Thr Leu Asp Ala Asn Gly Lys Val
260 265 270
       Thr Ala Leu Arg Ile Val Thr Tyr Leu Asn Glu Arg Asp Ser Lys Gly 275 280 285
20
       Ala Thr Val Glu Val Ala Leu Pro Arg Val Asp Asp Gly Thr Leu Pro
290 295 300
       Pro Pro Glu Phe Gly Pro Glu Leu Tyr Arg Leu Pro Leu Pro Asp Lys
305 310 315
       lle Leu Arg Asn His Trp Tyr Lys Tyr Glu Val Glu Ile
       (2) INFORMATION FOR SEQ ID NO:458
30
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 315 amino acids
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
35
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
40
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) MAME/KEY: misc feature
                   (B) LOCATION 1...315
45
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:458
       Met Ile Arg Thr Ile Leu Ser Arg Tyr Val Ser Ser Asn Phe Trp Ser
       Arg Gly Ala Thr Phe Phe Phe Thr Ile Phe Pro Ala Phe Ile Leu Ala
20 25 30
50
       Ala Thr Ala Leu Pro Ala Cys Gly Gly Gly Thr Ala Ser Gly Ser Asp
35 40
       Arg Thr Leu Ala Val Thr Ile Glu Pro Gln L's Tyr Phe Ile Glu Ser
50 60
55
       Ile Ala Asp Lys Ser Val Gln Val Val Ala Leu Val Pro Ala Gly Ser
65 70 75 80
       Asn Pro Glu Glu Tyr Asp Pro Ser Pro Thr Val Met Lys Arg Leu Ser
85 90 95
       Glu Ala Asp Ala Tyr Phe Tyr Ile Gly Gly Leu Gly Phe Glu Gln Arg
100 105 110
60
       Asn Leu Ala Ala Ile Arg Asp Asn Asn Pro L;s Leu Pro Leu Phe Glu
115 120 125
       Net Gly Lys Ala Leu Ala Asp Ala Gly Ser Ala Asp Leu His Gly Ser 130 $135$
65
        Cys Thr Asp His Ser His Thr Asp Leu His Ala His Asp Pro His Tyr
145 150 155 160
        Trp Ser Ser Val Val Gly Ala Lys Ala Leu Ser Arg Ala Ala Tyr Asp
165 170 175
        Ala Leu Val Glu Leu Tyr Pro Asn Glu Lys Asp Lys Trp Asp Lys Gly
180 190
 70
        His Asp Arg Leu Ash Gly Arg Ile Asp Ser Val Lys Arg Leu Val Asp
195 200 205
        Thr Het Phe Ala Asn Gly Lys Ala Asp Lys Ala Phe Val Ile Tyr His
 75
```

```
Pro Ser Leu Ser Phe Phe Ala Gln Glu Phe Gly Leu Arg Gln Ile Val
225 230 235 240
        Ile Glu Glu Asp Gly Lys Glu Pro Thr Ala Ala His Leu Arg Arg Val
245 250 255
        Ile Asp Gin Ala Arg Ala Asp Gly Val Arg Ile Val Phe Ile Gin Pro 260 265 270
        Glu Phe Glu Thr Arg Gln Ala Glu Asp Ile Ala Arg Glu Ile Gly Ala
275 280 285
        Arg Pro Val Arg Ile Asn Pro Leu Arg Ser Ser Trp Glu Glu Glu Ile
290 295 300
10
        Leu His Ile Ala Arg Ala Leu Ala His Glu Arg
                                 310
        (2) INFORMATION FOR SEQ ID NO:459
15
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 848 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
20
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
25
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAME/KEY: misc_feature (B) LOCATION 1...848
30
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:459
       Net Ile Gly Lys Lys Ile Phe Phe Ile Leu Leu Ala Leu Ile Ala Phe
35
       Ser Gly Leu Asn Ala Ala Thr Asp Thr Glu Phe Lys Tyr Pro Thr Asp 20 \\ 25 \\ 30
        Ala Asn Ile Ile Gly His Val Lys Asp Ser Lys Thr Gly Glu His Leu
35 40
40
        Val Gly Ile Thr Ile Ala Ile Lys Gly Thr Thr Phe Gly Thr Ser Thr 50 60
        Asp Ala Thr Gly His Tyr Leu Arg Asn Leu Arg Pro Gly Glu Ile
65 70 80
       Thr Leu Ile Het Arg Gly Het Gly Tyr Lys Ser Gln Glu Arg Val Val 85 90 95
45
        Arg Val Glu Lys Asp Lys Thr Ile Glu Val Asn Phe Glu Ala Glu Glu
100 105
       Asp Ala Ile Asn Leu Asp Glu Val Val Ile Ser Ala Asn Arg Glu Leu 115 120 125

Thr Leu Arg Arg Leu Ala Pro Thr Leu Val Asn Val Leu Asn Glu Lys 130 135 140
50

      Val Phe Ser Gln Val Asn Ala Ser Asn Leu Ala Gln Gly Leu Ser Phe

      145
      150
      155
      160

      Gln Pro Gly Val Arg Val Glu Asn Asn Cys Gln Asn Cys Gly Phe Asn
      165
      170
      175

55
        Gln Val Arg Ile Asn Gly Leu Asp Gly Arg Tyr Ala Gln Ile Leu Ile
180 185 190
        Asp Ser Arg Pro Ile Met Ser Ala Leu Ala Gly Val Tyr Gly Leu Glu
195 200 205
60
        Gln Ile Pro Ala Asn Met Ile Glu Arg Val Glu Val Val Arg Gly Gly
210 215 220
        Gly Ser Ala Leu Tyr Gly Ser Ser Ala Ile Ala Gly Val Val Asn Ile
225 230 235 240
       Ile Thr Lys Glu Pro Ser His Asn Ser Phe Thr Phe Asn Glu Ser Leu
245 250 255
65
       Ser Phe Thr Gly Phe Ser Lys Leu Asp Asn Asn Thr Asn Phe Asn Ala 260 265 270
       Ser Ile Val Ser Asp Asp Asp Arg Ala Gly Ala Met Val Phe Gly Gln 275 280 285
70
       Ala Arg Tyr Arg Asn His Trp Asp Ala Asn Asn Asp Gly Tyr Ser Glu
290 295 300
        Leu Gly Lys Ile Asp Ala Arg Ser Leu Gly Ala His Ser Tyr Leu Arg
305 310 315 320
                                                         315
        Leu Ser Asp Tyr Ser Lys Leu Thr Gly Glu Phe His Thr Ile Ser Glu
75
```

```
Phe Arg Arg Gly Gly Asp Arg Ile Asp Leu Pro Pro His Val Val Gly 340 345
       Val Ala Glu Gln Thr Asp His Ser Val Phe Ser Gly Asn Leu Lys Tyr
355 360 365
       Asp Leu Phe Ser Ser Asn Tyr Lys His His Phe Gln Ala Tyr Thr Ser
370 375 380
       Gly Gln Ile Val Asn Arg Lys Ser Tyr Tyr Gly Gly Ile Gly Glu Ile
385 390 395 400
       Asp Val Asn Gly His Pro Gly Gly Thr Glu Gly Tyr Pro Ile Pro Gln 405 415
10
       Asp Gln Tyr Gly Asn Asn Tyr Gly Val Thr Lys Gly Lys Thr Tyr Het 420 425 430
       Gly Gly Ile Gln Tyr Ser Tyr Asp Leu Asp Lys Phe Leu Leu Met Pro
       Ser Gln Leu Leu Phe Gly Ala Glu Tyr Thr Arg Asp Glu Leu Asn Asp
450 455 460
15
       Val Het Pro Ile Leu Ser Trp Gln Thr Gly Glu Asp Ala Asn Gly Asn
465 470 475 480
       Thr Ile Pro Leu Tyr Pro Glu Leu Asp Gln Asn Ile Asn Asn Tyr Ser
485 490 495
20
       Leu Phe Gly Gln Asn Glu Trp Lys Asn Asp Arg Trp Ser Ile Leu Val
500 505 510
       Gly Ala Arg Leu Asp Lys His Ser Glu Val Lys Asp Met Ile Leu Ser
515 520 525
       Pro Arg Thr Thr Leu Arg Phe Asn Val Asn Pro Asp Ile Asn Leu Arg
530 535 540
25
       Ala Thr Tyr Ala Lys Gly Phe Arg Ala Pro Gln Val Phe Arp Glu Asp
545 550 555 560
       Leu His Val Gly Val Val Gly Gly Glu Ala Gln Lys Val Phe Asn Asp
565 570 575
30
       Pro Asn Leu Lys Pro Glu Ile Ser His Ala Phe Ser Leu Ser Ala Asp
580 585 590
       Het Tyr His Arg Phe Gly Asn Val Gln Thr Asn Phe Leu Val Glu Gly
595 600 605
       Phe Tyr Thr Arg Leu Leu Asp Val Phe Thr Asn Glu Glu Gln Pro Asp
610 615 620
35
       Gln His Asp Gly Ile Lys Arg Tyr Thr Arg Ile Asn Gly Ser Gly Ala
625 630 635 640
       Lys Val Phe Gly Leu Asn Leu Glu Gly Lys Val Ala Tyr Lys Ser Phe
645 650 655
40
       Gln Leu Gln Ala Gly Leu Thr Leu Ala Ser Asn Lys Tyr Asp Glu Ala
660 665 670
       Gln Glu Trp Gly Leu Asn Thr Val Lys Asp Thr Asn Gly Ala Phe Val
675 680 685
       Thr Glu Ala Asn Ala Asn Gly Gln Gln Glu Tyr Lys Asn Glu Ser Het
690 695 700
45
       Thr Asp Thr Gln Ile Thr Arg Thr Pro Ser Val Tyr Gly Tyr Phe Thr
705 710 720
       Leu Ala Tyr Asn Pro Ala His Ser Trp Asn Ile Ala Leu Thr Gly Ala
725 736 735
50
       Tyr Thr Gly Gln Net Tyr Val Pro His Ala lle Glu Tyr Gly Val Lys
740 745 750
       Ser Ala Glu Leu Asp Ile Het Gln Asn Asn Pro Glu Ile Thr Asp Glu
755 760 765
       Thr Gly Lys Ala Pro Arg Ile Asp Glu Leu Lys Lys Thr Pro Ala Phe 770 780
55
       Phe Asp Leu Gly Leu Lys Val Gly Tyr Asp Phe His Val Phe Gln Ala
785 790 800
       Thr Glu Val Gln Leu Tyr Val Gly Net Asn Asn Ile Phe Asn Ser Phe
805 810 815
60
       Gin Lys Asp Phe Asp Arg Gly Ala Ala Arg Asp Ser Gly Tyr 11e Tyr
820 825 830
       Gly Pro Thr Gln Pro Arg Thr Gly Tyr Net Gly Leu Val Val Lys Phe
835 840 845
65
        (2) INFORMATION FOR SEQ ID NO: 460
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 202 amino acids
70
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) HOLECULE TYPE: protein
75
           (iii) HYPOTHETICAL: YES
```

```
(vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
 5
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...202
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:460
10
      The Leu Phe Ser Ser Pro Ser Leu Val Arg Ala Gln Ser Leu Phe Ser 20 25 30
15
      Thr Glu His Val Leu Gln Leu Tyr Asn Lys Ile Leu Tyr Gly Glu Ser 35 40 45
      Ala Ala Asp Thr Val Ala Glu Lys Thr Ala Gly Glu Ser Ala Phe Pro 50 60
      Phe Ile Asp Lys Leu Ile Asn Leu Gly Arg Thr Phe Leu Gly Lys Pro 65 70 75 80
20
      Tyr Arg Tyr Arg Gly Pro Ser Pro Trp Pro Het Asp Cys Ser Gly Tyr 85 90 95
      Val Ser Tyr Leu Tyr Ser Lys Phe Asp Ile Lys Leu Pro Arg Gly Ala
100 105 110
25
      Ala Ala Gln Ser Gln Tyr Thr Asn Pro Ile Glu Arg Glu Asp Val Arg
115 120 125
      Pro Gly Asp Leu Leu Phe Phe Lys Gly Arg Asn Ala Arg Ser Asn Arg
130 135 140
      Ile Gly His Val Ala Leu Val Val Ser Val Asp Glu Asp Asp Ile Thr
145 150 160
30
      Met Het His Ser Arg Asn Ser Arg Gly Ile Val Ile Glu Lys Leu Asn 165 170 175
      Arg Ser Ala Tyr Phe Ser Arg Arg Leu Val Ser Tyr Gly Arg Val Pro
180 185 190
35
      Gly Ala Lys Arg Val Ile Pro Arg Lys Ser
      (2) INFORMATION FOR SEQ ID NO:461
40
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 455 amino acids
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
45
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
50
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...455
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:461
      Met Lys Arg Thr Ile Leu Leu Thr Ala Leu Thr Val Leu Ser Ser Leu
                                            10
                                                                1.5
60
      Ser Leu Leu Arg Ala Gln Asn Glu Ser Glu Ala Ser Thr Asn Pro Met
                                       25
      Ser Gly Leu Ser Leu Glu Asp Cys Ile Arg Ile Ala Lys Glu Arg Asn
              35
                                   40
      Leu Asn Leu Arg Arg Gln Glu Ile Glu Gln Glu Asn Arg Ile Ile Ser
50 55 60
65
      Leu Asp Ala Ala Arg His Ser Phe Leu Pro Ser Val Asn Ala Gly Ile
65 70 75 80
      Gly His Asn Tyr Ser Phe Gly Arg Ser Lys Asp Lys Thr Gly Val Thr
85 90 95
70
      Val Asp Arg Ser Ser Met Asn Thr Asn Leu Ser Ile Gly Ala Ser Val
100 105 110
      Glu Val Phe Ser Gly Thr Arg Arg Leu His Asp Leu Lys Gln Gln Lys 115 120 125
      Tyr Asn Val Glu Asp Gly Ile Ala Arg Leu Gln Lys Ala Arg Glu Asp
75
```

```
Leu Ser Leu Gln Ile Ala Ala Leu Tyr Ile Asn Leu Leu Phe Arg Gln
145 150 155 160
       Glu Het Thr Arg Thr Ala Glu Thr Gln Leu Ala Leu Ile Arg Glu Gln
165 170 175
       Arg Asn Arg Thr Ala Glu Met Val Arg Val Gly Lys Trp Ala Glu Gly
180 185
       Lys Leu Leu Asp Ile Asn Ala Gln Met Ala Lys Asp Glu Gln Leu Leu
195 200 205
       Val Gln Tyr Arg Ser Glu Glu Glu Leu Ala Arg Leu Asp Leu Gly Gln
210 220
10
       Ala Leu Glu Leu Glu His Pro Glu Ser Ile Ala Val Lys Ala Pro Asp
225 230 235 240
       Thr Asp Val Leu Val Ala Glu Arg Leu Gly Ser Leu Leu Ala Pro Glu
245 250 255
       Glu Ile Tyr Arg Thr Ala Leu Gly Leu Lys Pro Ala Leu His Ser Ser
260 265 270
15
       Glu Leu Gln Ile Ala Ser Ala Arg Glu Gly Leu Ala Ser Ala Arg Ala
275 280 285
       Ala Tyr Phe Pro Thr Leu Ser Leu Ser Ala Gly Tyr Ser Asn Gly Tyr
290 295 300
20
       Phe Arg Asp Leu Gly Lys Glu Tyr Ala Ala Ile Asn Pro Ser Phe Ser
305 310 320
       Glu Gln Trp Lys Asn Asn Gly Ser Tyr Ser Ile Gly Leu Ser Leu Asn 325 330 335
       Ile Pro Ile Phe Ser Ala Met Gln Thr Gln Asp Arg Val Arg Ser Ser 340 345 350
25
                    340
       Arg Leu Gln Tie Arg Ser Ser Glu Leu Arg Leu Val Glu Glu Lys Lys
355 360 365
       Ala Leu Tyr Lys Glu fle Arg Gln Ala Tyr Ser Asn Ala Val Ala Ala
370 375 380
30
          370
       Asp Lys Ala Ile Ala Ala Ala Glu Asn Ser Lys Ala Ala Thr Leu Lys
385 390 395 400
       Ala Tyr Glu Tyr Ala Arg Asp Ser Phe Glu Ala Gly Arg Leu Ser Ala
405 410 415
       Tyr Glu Tyr Ala Glu Ala Lys Thr Lys Tyr Ala Leu Ser Gln Val Glu
420 430
35
       Glu Leu Arg Ala Lys Tyr Asp Phe Ile Tyr Lys Ala Lys Val Leu Asp
435 440 445
       Phe Tyr Gln Gly Lys Asp Phe
40
       (2) INFORMATION FOR SEQ ID NO: 462
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 444 amino acids (B) TYPE: amino acid
45
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
50
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...444
60
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:462
       Het Arg Phe Gln His Tyr Leu Ile Cys Thr Ala Ala Val Ala Ala Leu
                                               10
       Ala Ala Ash Pro Leu Thr Gly Gln Ser Ash Het Thr Leu Glu Glu Cys
65
                                           25
       Ile Asp Tyr Ala Arg Arg His Ser Ser Ala Val Ala Leu Ser Ala Ala 35 40 45
       Glu Leu Glu Gln Ser Lys Ala Asp Tyr Leu Gln Ala Val Gly Asn Phe
50 60
       Leu Pro Arg Val Ser Ala Gly Thr Gly Ala Ser Trp Asn Phe Gly Arg 65 70 75
70
       Gly Leu Asp Ala Glu Thr Asn Thr Tyr Thr Asp Ile Asn Ser Phe Asn
85 90
        Asn Ser Tyr Ser Ile His Ala Thr Met Thr Leu Phe Asp Gly Leu Gln
75
```

```
Ser Val Tyr Arg Leu Arg Met Ala His Ala Arg Arg Glu Ala Ser Arg 115 $120 $125
       Leu Ser Val Arg Glu Gln Gln Glu Leu Ala Ala Leu Gly Thr Thr Glu
                                  135
                                                       140
           130
       Ala Tyr Tyr Asp Leu Val Tyr Ala Arg Gln Net Gln Glu Leu Ala Met
145 150 160
       Gln Lys Tyr Glu Glu Ser Ser Arg Leu His Arg Gln Thr Ala Arg Het
165 170 175
       Glu Glu Leu Gly Met Lys Ser Arg Pro Asp Val Leu Glu Met Gln Ser
180 185 190
10
       Arg Met Ala Gly Asp Arg Leu Ala Leu Thr Gln Ala Asp Asn Gln Cys
195 200 205
       Ile Ile Ala Leu Ile Arg Leu Lys Glu Lys Het Asn Phe Pro Ile Asp
210 215 220
       Asp Glu Leu Val Val Asp Asp Met Pro Ala Asp Ser Leu Ser Ala Asp
225 230 235 240
15
       Met Ala Glu Ser Asp Ser Ser Ala Gly Val Phe Ala Arg Ala Ala His
245 250 255
       His His Pro Val Leu Leu Arg Ala Lys Leu Asp Glu Gln Ala Ala Thr
260 265 270
20
       Asp Arg Leu Arg Ala Ala Arg Gly Ala Phe Leu Pro Ser Val Ser Val 275 280 285
       Ser Gly Gly Trp Asn Thr Gly Phe Ser Arg Phe Leu Asn Gly Ser Asp
290 295 300
       Tyr Thr Pro Phe Ser Glu Gln Phe Arg Asn Arg Arg Gly Glu Tyr Val
305 310 315 320
25
       Ser Leu Asn Leu Ser Ile Pro Ile Phe Ser Gly Phe Ser Leu Val Ser 325 330 335
       His Leu Arg Gln Ala Arg Ala Glu Arg Arg Ala Ala Ile Val Arg Arg 340 $340$
30
       Gly Glu Ala Glu Arg Arg Leu Tyr Ser Glu Ile Ala Gln Ala Net Ala
355 360 365
       Asp Arg Asp Ala Ala Leu Ala Ser Tyr Arg Gln Ala Lys Glu His Thr 370 \hspace{1cm} 375 \hspace{1cm} 380
35
       Asp Ala Net Gln Thr Ala Tyr Glu Ala Val Leu Gln Arg Tyr Glu Glu 385 390 395 400
       Gly Leu Asn Thr Ala Ile Asp Leu Thr Thr Gln Ala Asn Arg Leu Leu
405 410 415
       Asp Ala Arg Val Gln Arg Leu Arg Ala Ala Net Thr Tyr Arg Leu Lys
420 425 430
40
       Cys Lys Leu Ile Ala Tyr Tyr Gly Cys Leu Ser Asp
       (2) INFORMATION FOR SEQ ID NO:463
45
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 940 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
50
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KET: misc_feature
60
                  (B) LOCATION 1...940
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:463
       Met Asn Lys Phe Tyr Lys Ser Leu Leu Gln Ser Gly Leu Ala Ala Phe
65
       Val Ser Het Ala Thr Ala Leu Thr Ala Ser Ala Gln Ile Ser Phe Gly 20 25 30
       Gly Glu Pro Leu Ser Phe Ser Ser Arg Ser Ala Gly Thr His Ser Phe 35 40 45
70
       Asp Asp Ala Met Thr Ile Arg Leu Thr Pro Asp Phe Asn Pro Glu Asp 50 55 60
       Leu Ile Ala Gln Ser Arg Trp Gln Ser Gln Arg Asp Gly Arg Pro Val
65 75 80
       Arg Ile Gly Gln Val Ile Pro Val Asp Val Asp Phe Ala Ser Lys Ala
75
```

| | Ser | His | Ile | Ser 100 | Ser | Ile | Gly | Asp | Vā1 105 | Asp | Val | Tyr | Arg | Leu 110 | Gln | Phe |
|----|------------|---------------------|--------------------|------------|------------|-------------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|
| | Lys | Leu | Glu 115 | | Ala | ГЛЗ | Ala | Ile 120 | | Leu | Tyr | Tyr | Asp 125 | | Phe | Asn |
| 5 | Ile | Pro 130 | | Gly | Gly | Arg | Leu 135 | | Ile | Туг | Thr | Pro 140 | | His | Glu | Ile |
| | Val 145 | Leu | Gly | Ala | Tyr | Thr 150 | Asn | Ala | Thr | His | Arg 155 | Arg | Asn | Gl7. | Ala | Phe 160 |
| 10 | Ala | Thr | Glu | Pro | Val 165 | | Gly | Ser | Glu | Leu 170 | Ile | Met | Asp | Tyr | Glu 175 | Val |
| | Ser | Arg | G1 λ. | Gly 180 | Thr | Leu | Pro | Asp | Ile 185 | Lys | Ile | Ser | Gly. | Ala 190 | Glү | T;r |
| | | | 195 | | Val | | | 200 | | | | | 205 | | | |
| 15 | Ile | Gly 210 | Glu | Asp | Asp | Ser | Asp 215 | Ser | Asp | CAR | Glu | 11e 220 | Asn | Ile | Asn | Cys |
| | Pro 225 | Glu | Gly | Ala | Asp | Trp 230 | Gln | Alā | Glu | Lys | Asn 235 | G1?. | Val | Val | Gln | Het 240 |
| 20 | Ile | Het | Val | L'æ | Gly 245 | Gln | Туr | Ile | Ser | Het 250 | СЛа | Ser | Gly | Asn | Leu 255 | Leu |
| | Asn | Asn | Thr | L;s 260 | G1 ?. | Asp | Phe | Thr | Pro 265 | Leu | Ile | Il€ | Ser | Ala 270 | 61 y | His |
| | Cys | Ala | Ser 275 | Ile | Thr | Thr | Asn | Phe 280 | Gl) | Val | Thr | Gln | Ser 285 | Glu | Leu | Asp |
| 25 | | 290 | | | Thr | | 295 | | | | | 300 | | | | |
| | 305 | | | | Phe | 310 | | | | | 315 | | | | | 320 |
| 30 | Ala | Phe | Leu | Pro | 11e 325 | Lys | Gly | Lys | Ser | Asp 330 | Gly | Leu | Leu | Leu | Gln 335 | Leu |
| | Asn | Asp | Glu | Val 340 | Pro | Leu | Arg | Туг | Arg 345 | Val | Tyr | Tyr | Asn | G1 y 350 | Trp | Asp |
| | Ser | Thr | Pro 355 | Asp | Ile | Pro | Ser | Ser 360 | Gl; | Ala | Gly | Ile | His 365 | His | Pro | Ala |
| 35 | Gly | А <i>г</i> р 370 | Ala | Het | Lys | Ile | Ser 375 | Ile | Leu | Lys | Lys | Thr 380 | Pro | Ala | Leu | Asn |
| | Thr 385 | Trp | Ile | Ser | Ser | Ser 390 | Gly | Ser | Glγ | Gly | Thr 395 | Asp | Asp | His | Fhe | Tyr 400 |
| 40 | | | | | Gln 405 | | | | | 410 | | | | | 415 | |
| | Leu | Phe | Asn | Gln 420 | Asn | Lys | His | Val | Val 425 | Gly | Thr | Leu | Thr | Gly 430 | Gly | Ala |
| | Gly | Asn | Су <i>в</i> 435 | G17. | Gly | Thr | Glu | Phe 440 | Tyr | Gly | Arg | Leu | Asn 445 | Ser | His | Trp |
| 45 | Asn | Glu 450 | Tyr | Ala | Ser | Asp | Gly 455 | Asn | Thr | Ser | Arg | Met 460 | Asp | Ile | Tyr | Leu |
| | Asp 465 | Pro | Gln | Asn | Asn | Gl ₃ ' | | Thr | Thr | Ile | Leu 475 | Asn | Gly | Thr | Tyr | Arg 480 |
| 50 | Asp | Gly | Tyr | Γ.'.ε | Pro 485 | Leu | Pro | Ser | Val | Pro 490 | Arg | Leu | Leu | Leu | Gln 495 | Ser |
| | Thr | Gly | Asp | Gln 500 | Val | Glu | Leu | Asn | Trp 505 | Thr | Ala | Val | Pro | Ala 510 | Asp | Gln |
| | Ty.r | Pro | Ser 515 | Ser | Туr | Gln | Val | Glu 520 | | His | Ile | Phe | Arg 525 | Asn | Gly | L∵s |
| 55 | Glu | 11e 530 | | Thr | Thr | Lys | Glu 535 | Leu | Ser | Туr | Ser | Asp 540 | Ala | Ile | Asp | Glu |
| | 545 | | | | Ser | 550 | | | | | 555 | | | | | 560 |
| 60 | | = | | | Pro 565 | | | _ | | 570 | | _ | | | 575 | |
| | _ | | | 580 | | | | | 585 | - | | | | 590 | | |
| | | _ | 595 | | Pro | | | 600 | _ | _ | | | 605 | | | |
| 65 | | 610 | | | Ser | | 615 | | | _ | | 620 | | | | |
| | 625 | | | - | Thr | 630 | | | | _ | 635 | | | | | 640 |
| 70 | | | | | Pro 645 | | | | | 650 | | | | | 655 | |
| | | _ | | 660 | | | _ | | 665 | | | | | 670 | | |
| | Pro | Ser | Ala 675 | | Asp | Ser | Thr | Phe 680 | | Leu | Phe | Leu | Lys 685 | | Asn | Thr |
| 75 | Asn | Arg | Arg | Leu | Gln | Lys | Val | | | Pro | Ser | Asp | Trp | Gln | Ala | Gl; |

```
695
                                                            700
       Thr Trp Leu Arg Ile Asn Leu Asp Lys Pro Phe Pro Val Asn Asn Asp 705 710 715 720
       His Met Leu Phe Ala Gly Ile Arg Met Pro Asn Lys Tyr Lys Leu Asn 725 730 735
       Arg Ala Ile Arg Tyr Val Arg Asn Pro Asp Asn Leu Phe Ser Ile Thr 740 745 750
       Gly Lys Lys 11e Ser Tyr Asn Asn Gly Val Ser Phe Glu Gly Tyr Gly 755 \hspace{1.5cm} 760 \hspace{1.5cm} 765
10
       Ile Pro Ser Leu Leu Gly Tyr Met Ala Ile Lys Tyr Leu Val Val Asn 770 \phantom{000}775\phantom{000}780\phantom{000}
       Thr Asp Ala Pro Lys Ile Asp Net Ser Leu Val Gln Glu Pro Tyr Ala
785 790 795 800
       Lys Gly Thr Asn Val Ala Pro Phe Pro Glu Leu Val Gly Ile Tyr Val
805 810
15
                        805
       Tyr Lys Asn Gly Thr Phe Ile Gly Thr Gln Asp Pro Ser Val Thr Thr 820 825 830
       Tyr Ser Val Ser Asp Gly Thr Glu Ser Asp Glu Tyr Glu Ile Lys Leu
835 840 845
20
       Val Tyr Lys Gly Ser Gly Ile Ser Asn Gly Val Ala Gln Ile Glu Asn
850 855 860
       Asn Asn Ala Val Val Ala Tyr Pro Ser Val Val Thr Asp Arg Phe Ser 865 870 875
       lle Lys Asn Ala His Net Val His Ala Ala Ala Leu Tyr Ser Leu Asp
885 890 895
25
       Gly Lys Gln Val Arg Ser Trp Asn Asn Leu Arg Asn Gly Val Thr Phe
900 905 910
       Ser Val Gln Gly Leu Thr Ala Gly Thr Tyr Het Leu Val Met Gln Thr
915 920 925
30
       Ala Asn Gly Pro Val Ser Gln Lys Ile Val Lys Gln
        (2) INFORMATION FOR SEQ ID NO:464
35
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 670 amino acids(B) TYPE: amino acid
                   (D) TOPOLOGY: linear
40
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
45
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...570
50
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:464
       Het Lys Tyr Leu Ile Arg Leu Phe Leu Ser Leu Het Leu Leu Ser Leu
1 5 15
                                                 10
55
       Trp Thr Gly Cys Thr His Glu Glu Leu Ser Ile Cys Asp Gly Glu Asn 20 \\ 25 \\ 30
       Thr Leu Val Leu Arg Val Glu Thr Gly Lys Ala Pro Asn Ala Arg Ala 35 40 45
       Thr Glu Pro Gly Gln Gly Ile Tyr Asn Glu Asn Lys Val Gly Ser Ile
50 55
60
       Ser Val Leu Phe Tyr Leu Glu Gly Gln Leu Arg Trp Gln Val Lys Ser 65 70 75
       Thr Asp Tyr Gln Ile His Glu Gly Ala Tyr Ile Ile Pro Val Lys Glu
85 90 95
65
       Gln Net Arg Pro Leu Phe Asn Gly Asn Asn Asn Phe Ser Tle Tyr Val
100 105 110
       Val Ala Asn Leu Asp Phe Asn Ala Pro Ala Thr Glu Ala Ala Leu Ser
115 120 125
       Gln Fhe Val Val Glu Lys Ser Ile Glu Val Ser Ser Thr Thr Ala Pro
130 135 140
70
       130 135 140

Ala Asp Phe Val Met Leu Ala His Gly Asn Lys Gln Ile Asn Met Ala
145 150 155 160
       Thr Thr Glu Gly Lys Leu Leu Gly Asp Tyr Lys Leu Lys Arg Val Ala
165 170 175
75
       Ala Lys Ile Arg Het Ile Lys Pro Thr Ile Asn Val Gln Gly Tyr Glu
```

```
190
                                             185
      Val Val Gly Asn Ile Gln Ala Lys Phe Arg Asn Ser Val Thr Lys Gly
195 200 205
      Phe Leu Thr Thr Glu Ala Gln Glu Ile Pro Ala Ala Ala Ser Tyr Lys
210 215 220
 5
      Thr Ser Glu Tyr Leu Asp Ile Ala Glu Ser Ala Pro Ala Asn Ser Ile
225 230 235 240
       His Phe Tyr Ser Tyr Tyr Asn Lys Trp Thr Leu Ser Thr Pro Glu Lys 245 250 255
      Arg Pro Glu Phe Phe Ile Met Val Lys Phe Lys Lys Thr Gly Gln Pro 260 265 270
10
       Asp Asn Thr Ala Lys Pro Tyr Tyr Tyr Arg Val Pro Leu Glu Ser Gln 275 280 285
       Asp Asn Gln Val Lys Ser Asn Val Leu Tyr Asn Leu Asn Val Lys Ile
290 295 300
15
       Glu Île Leu Gly Ser Leu Gln Glu Pro Glu Ala Val Ser Val Asn Gl;
305 310 315 320
       Thr Leu Ala Ile Glu Glu Trp Ile Leu His Gln Asp Ala Phe Asn Leu
325 330 335
       Pro Ala Thr Asn Tyr Leu Ile Val Glu Gln His Glu Ile Phe Met Asn 340 350
20
       Asn Val Asn Thr Tyr Ser Val Lys Tyr Gln Thr Ser Gln Lys Pro Ile
355 360 365
       Ser Ile Ser Ile Gln Ser Val Thr Phe Ser Tyr Val Ser Ser Asp Gly 370 375
25
       Thr Gln His Asn Asp Leu Val Ala Ser Ser Ser Asp Gln Tyr Pro Thr 385 390 395 400
       Ile Thr Ser Asp Asn Thr Ser Ile Ile Ile Thr Ser Lys Ile Pro Val
405 410 415
       Asn Asn Val Pro Lys Lys Ile Val Phe Glu Val Thr Asn Gly Val Ala 420 430
30
       Gly Leu Lys Glu Thr Val Thr Val Leu Gln Tyr Pro Ala Gln Phe Ile
435 440 445
       Val Asn Thr Leu Gly Thr Ala Ser Ala Trp Arg Pro Asp Gly Ser Leu
450 455 460
35
       Ala Pro Gly Leu Asn Asn Lys Ala Ile Tyr His Val Val Val Leu Val
465 470 475 480
       Pro Pro Glu Asn Leu Phe Glu Asp Gly Thr Gln Thr Ile Ile Gly Tyr
485 490 495
       Pro Pro Thr Glu Thr Ile Ser Phe His Lys Lys Glu Asn Asn Thr Tyr 500 510
40
       Pro Ile Val Trp Ser Asp Thr Asn Thr Thr Lys Gln Asp Leu Glu Thr
515 520 525
       Ser Arg Het Ile Ser Pro Ser Phe Glu Leu Ala Ser Gln Leu Gly Ala
530 535 540
45
       Thr Leu Pro Het Pro Tyr Leu Glu Tyr Trp Pro Gly Thr Ser Tyr Leu 545 550 555 560
       Leu Asp Tyr Ser Gly Asn Tyr Asn Asn Lys Arg Tyr Ala Leu Phe Asn 565 570 575

Cys Ala Phe Tyr Trp Glu Lys Arg Lys Val Asn Asn Glu Glu Ile Lys 580 590
50
       Phe Asp Asp Trp Arg Leu Pro Thr Glu Ala Glu Ile Lys Leu Ile Asp
595 600 605
       Lys Leu Gln His Asn Glu Gln Ser Ala Val Gln Ala Île Het Thr Gly
610 615 620
55
       Asn Tyr Tyr Trp Asp Ser Tyr Ser Ala Asn Gly Ser Tyr Lys Het Gln 625 630 635
       Gly Gly Gly Gln Gly Asn Ser Ser Lys Ala Tyr Val Arg Cys Val
645 650 655
       Arg Asp Val Lys Lys Pro Ile Arg Asp Lys Lys Ser Gly Lys
60
                      660
        (2) INFORMATION FOR SEQ ID NO: 465
65
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 1282 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
 70
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
 75
                    (A) ORGANISH: Porphyromonas gingivalis
```

```
(ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...1282
 5
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 465
       Met Arg Lys Ile Leu Ser Phe Leu Met Met Cys Ser Leu His Leu Gly
                                                 10
10
       Leu Gln Ser Gln Thr Trp His Gly Asp Pro Asp Ser Val Ala Leu
20 25 30
       Pro Ser Ile Gly Ile Gln Glu Ser Ser Cys Thr Arg Ile Thr Phe Glu
35 40
       Val Val Phe Pro Gly Phe Tyr Ser Val Glu Lys Arg Glu Gly Asn Gln 50 60
15
       Val Phe Gln Arg Ile Ser Het Pro Gly Cys Gly Ser Phe Gly Asn Leu
65 70 75 80
       Gly Glu Ala Glu Leu Pro Val Leu Lys Lys Het Ile Ala Val Pro Glu 85 \ 90 \ 95
20
       Fhe Ser Thr Ala Asn Val Ala Val Lys Ile Lys Glu Thr Glu Thr Phe 100 \hspace{1cm} 105 \hspace{1cm} 110
       Asp Asn Tyr Asn Ile Tyr Pro Asn Pro Thr Tyr Val Val Glu Glu Leu
115 120 125
       Pro Glu Gly Gly Thr Tyr Leu Val Glu Ala Phe Ala Ile Asn Asn Asp
130 135
25
       Tyr Tyr Ser Gln Asn Val Ser Leu Pro Ser Thr His Tyr Val Tyr Ser
145 150 155 160
       Gln Asp Gly Tyr Phe Arg Ser Gln Arg Phe Ile Glu Val Thr Leu Tyr 165 170 175
30
       Pro Phe Arg Tyr Asn Pro Val Arg Gln Glu Ile Leu Phe Ala Lys Lys
180 185 190
       Ile Glu Val Thr Ile Thr Phe Asp Asn Pro Gln Pro Pro Leu Gln Lys
195 200 205
       Asn Thr Gly Ile Phe Asn Lys Val Ala Ser Ser Ala Phe Ile Asn Tyr
210 215 220
35
       Glu Ala Asp Gly Lys Ser Ala Ile Glu Asn Asp Met Val Phe Ser Arg
225 230 235 240
       Gly Thr Thr Tyr Ile Ser Gly Asn Val Ala Ser Asn Leu Pro Gln
245 250 255
40
       Asn Cys Asp Tyr Leu Val Ile Tyr Asp Asp Het Phe Asn Val Asn Gln
260 265 270
       Gln Pro His Asp Glu Ile Lys Arg Leu Cys Glu His Arg Ala Phe Tyr
275 280 285
       Asn Gly Phe Asp Val Ala Ala Val Ser Ile Lys Asp Val Leu Asn Ser
290 295 300
45
       Phe Pro Ser Asn Ala Thr Ser Tyr Ile Asn Glu Thr Lys Leu Lys Asn 305 310 315 320
       Phe Ile Arg Ser Val Tyr Asn Gln Ser Asn Ala Lys Arg Thr Leu Asp
325 330 335
50
       Giy Lys Leu Gly Tyr Val Leu Leu Ile Gly Lys Pro Leu Ser Lys Tyr 340 345 350
       Leu Ala Asp Thr Asp Asn Thr Lys Val Pro Thr Ser Phe Ile His Asn 355 360 365
       Val Ser Leu Ile Pro Ser His Pro Thr Phe Gly Ser Ile Cys Ala Ser
370 375 380
55
       Asp Tyr Phe Phe Ser Cys Val Ser Pro Leu Asp Thr Val Gly Asp Leu
385 390 395
       Phe Ile Gly Arg Phe Ser Val Thr Asn Ala His Glu Leu His Asn Leu
405 415
60
       Ile Glu Lys Thr Ile Asn Lys Glu Ile Ser Tyr Asn Fro Ile Ala His
420 425 430
       Lys Asn Ile Leu Tyr Ala Glu Gly Lys Gly Cys Asp Ala Pro Ile Leu
435 440 445
       Arg Leu Phe Leu Lys Glu Ile Ala Ser Gly Tyr Thr Val Asn Ser Ile
450 455 460
65
                                455
       Leu Lys Ser Asn Gln Val Ser Ala Ile Asp Ser Ile Phe Asp Cys Leu
465 470 475
       Asn Asn Gly Ser His His Phe Tyr Phe Asn Thr His Gly Met Pro Thr 485 490 495
70
       Val Trp Gl; Ile Gly Gln Gly Leu Asp Val Asn Thr Leu Thr Ala Arg
500 510
       Leu Asn Asn Thr Ser Ser Gln Gly Leu Cys Thr Ser Leu Ser Cys Ser 515 520 525
       Ser Ala Val Ala Asp Ser Thr Ile Arg Ser Leu Gly Glu Val Leu Thr
75
                                   535
```

| | | туг | Ala | Pro | Asn | | G) ۲ | Phe | Ser | Alā | Phe | Leu | Gl y | Gly | Ser | Arg 560 |
|------------|------------|------------|------------|------------|------------|------------|-----------|-----------|------------|-----------------|------------|------------|------------|------------|------------|-------------|
| | 545 Ala | Thr | Gln | туг | Ala 565 | 550 Val | Туг | Leu | Glu | Gly 570 | 555 Pro | Сув | Pro | Pro | Ser 575 | |
| 5 | Phe | туг | Glu | Tyr 580 | | Pro | Туг | Ser | Leu 585 | | His | Asn | Leu | Ser 590 | _ | Vāl |
| | | | Glu 595 | | | | | 600 | | | | | 605 | | | |
| 10 | | 610 | Ser | | | | 615 | | | | | 620 | | | | |
| | 625 | | Ala | | | 630 | | | | | 635 | | | | | 640 |
| | | | Ile | | 645 | | | | | 650 | | | | | 655 | |
| 15 | | | Pro | 660 | | | | | 665 | | | | | 670 | | |
| | | | Het 675 | | | | | 680 | | | | | 685 | | | |
| 20 | | 690 | Glu | | | | 695 | | | | | 7,00 | | | | |
| | 705 | | Ala | | | 71.0 | | | | | 715 | | | | | 720 |
| 25 | | | Phe Phe | | 725 | | | | | 730 | | | | | 735 | |
| 25 | | | Asn | 740 | | | | | 745 | | | | | 750 | | |
| | | | 755 Leu | | | | | 760 | | | | | 765 | | | |
| 30 | | 770 | | | | | 775 | | | | | 780 | | | | |
| | 785 | | Ile | | | 790 | | | | | 795 | | | | | 800 |
| 35 | | | Thr | | 805 | | | | | 810 | | | | | 815 | |
| 33 | | | Asp | 820 | | | | | 825 | | | | | 830 | | |
| | | | 835 Asp | | | | | 840 | | | | | 845 | | | |
| 40 | | 850 | | | | | 855 | | | | | 860 | | | | |
| | 865 | | Leu | | | 870 | | | | | 875 | | | | | 880 |
| 45 | | | His | | 885 | | | | | 890 | | | | | 895 | |
| | | | Ala | 900 | | | | | 905 | | | | | 910 | | |
| | _ | | 915 Thr | | | | | 920 | | | | | 925 | | | |
| 50 | | 930 | | - | | | 935 | | | | | 940 | | | | |
| | 945 | | Thr | | | 950 | | | | | 955 | | | | | 960 |
| 55 | | | Asn | | 965 | | | Pro | Asn | 970 Gln | Val | | | Thr | 975 Pro | |
| | Leu | Phe | Ile | | | Pro | Phe | Trp | Asp | Gl ₂ | | Pro | | | | Ser |
| 60 | Gly | | 995 Ser | | Ala | Glu | | | | Phe | Gln | | | | Asp | C7.2 |
| 60 | | _ | 0 Asn | Ser | Vsb | | | | Ala | Lys | | Ala | 020 Leu | Lys | Met | Met 1040 |
| | 102 Val | | Thr | туг | | | | Asp | Phe | | Ile | 035 Ala | Ala | Leu | | |
| 65 | Leu | Phe | Arg | Ile 106 | | | Net | Ser | Gl y 10 | | | Туr | Glu | | | |
| | Asp | Tyr | Phe 107 | Arg | | Asn | Pro | Thr 10 | Ile | | Ser | Ser | | _ | | Phe |
| 70 | Pro | Thr 109 | Ala | | Phe | Leu | Ser 10 | Ala | | Сув | Asp | Ile | | | Glu | Asn |
| , , | Туг 110 | Gln | Ser | Ala | Ile | Asp 11 | Trp | | Glu | Asn | | | | Ser | Glu | Il∈ 1120 |
| | | | Gln | Asp | Ser 112 | Val | | Ala | Val | | | | G1 y | Asp | | |
| <i>7</i> 5 | Trp | Asn | Het | Gln | | | Ser | Leu | Arg | | | Gl y | Ile | Asp | | |

```
1140
                                            1145
       Ile Leu Ser Cys Glu Gln Arg Lys Ser Leu Glu Ser His Gln Asn Val
1155 1160 1165
       Lys Asn Tyr Leu Leu Ser Thr Leu Pro Glu Ser Thr Gly Thr Leu Leu 1170 1180
 5
       Pro Pro Leu Glu Cys Asn Lys Ser Ser Leu Asp Lys Ser Lys Ile Ile
1185 1190 1195
       Ser Ile Ser Pro Asn Pro Ala Lys Ala Val Val Thr Ile Ile Tyr Tyr
1205 1210 1215
10
       The Asp Asn Pro Ser Cys Ser Val Ile Lys Ile Tyr Gly Ile Asn Gly 1220 1225 1230
       Ala Ser Ala Asp Ile Thr Gly Leu Pro Lys His Leu Ser Glu Gly Tyr
1235 1240 1245
                                     1240
                                                  1245
       Tyr Ser Ile Gln Fhe Asn Thr Ser Asn Phe Asp Pro Gly Fhe Tyr Leu
15
       1250 1255 1260
Val Thr Leu Asn Val Asp Gln Lys Ile Ile Asp Thr Glu Lys Leu Arg
       1265
       Ile Ly\varepsilon
20
       (2) INFORMATION FOR SEQ ID NO: 466
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1274 amino acids
25
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
30
          (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
35
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...1274
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 466
40
       Het Met Cys Ser Leu His Leu Gly Leu Gln Ser Gln Thr Trp His Gly
       Asp Pro Asp Ser Val Ala Ala Leu Pro Ser Ile Gly Ile Gln Glu Ser 20 25 30
45
       Ser Cys Thr Arg Ile Thr Phe Glu Val Val Fhe Pro Gly Phe Tyr Ser 35 40 45
       Val Glu Lys Arg Glu Gly Asn Gln Val Phe Gln Arg Ile Ser Met Pro
50 55 60
       Gly Cys Gly Ser Phe Gly Asn Leu Gly Glu Ala Glu Leu Pro Val Leu 65 75 80
50
       Lys Lys Het Ile Ala Val Pro Glu Phe Ser Thr Ala As<br/>n Val Ala Val 85 90 95
       Lys Ile Lys Glu Thr Glu Thr Phe Asp Asn Tyr Asn Ile Tyr Pro Asn 100 105 110
       Pro Thr Tyr Val Val Glu Glu Leu Pro Glu Gly Gly Thr Tyr Leu Val 115 120 125
55
       Glu Ala Phe Ala Ile Asn Asn Asp Tyr Tyr Ser Gln Asn Val Ser Leu
130 135
       Pro Ser Thr His Tyr Val Tyr Ser Gln Asp Gly Tyr Phe Arg Ser Gln
145 150 160
60
       Arg Phe Ile Glu Val Thr Leu Tyr Pro Phe Arg Tyr Asn Pro Val Arg
165 170 175
       Gln Glu Ile Leu Phe Ala Lys Lys Ile Glu Val Thr Ile Thr Phe Asp
180 185 190
65
       Asn Pro Gln Pro Pro Leu Gln Lys Asn Thr Gly Ile Phe Asn Lys Val
195 200 205
       Ala Ser Ser Ala Phe Ile Asn Tyr Glu Ala Asp Gly Lys Ser Ala Ile 210 215 220
       Glu Asn Asp Net Val Phe Ser Arg Gly Thr Thr Thr Tyr Ile Ser Gly 225 230 235 240
70
                                                235
       Asn Val Ala Ser Asn Leu Pro Gln Asn Cys Asp Tyr Leu Val Ile Tyr
245 250 255
       Asp Asp Het Phe Asn Val Asn Gln Gln Pro His Asp Glu Ile Lys Arg
                    260
                                          265
75
       Leu Cys Glu His Arg Ala Phe Tyr Asn Gly Fhe Asp Val Ala Ala Val
```

| | | | 275 | | | | | 280 | | | | | 285 | | | |
|------------|------------|------------|------------|-----|------------|------------|------------|------------|-----|-----|------------|------------|------------|-----|-----|------------|
| | Ser | 11e 290 | r?.e | Asp | Val | Leu | Asn 295 | Ser | Phe | Pro | Ser | Asn 300 | Ala | Thr | Ser | Tyr |
| 5 | Ile 305 | | Glu | Thr | Lys | Leu 310 | | Asn | Phe | Ile | Arg 315 | | Val | Туг | Asn | Gln 320 |
| Ü | Ser | | | | Arg 325 | Thr | | | | 330 | | | | | 335 | |
| | | | | 340 | Leu | | | | 345 | | | | | 350 | | |
| 10 | | | 355 | | Phe | | * | 360 | | | | | 365 | | | |
| | | 370 | | | Ile | | 375 | | | | | 380 | | | | |
| 15 | 385 | | | | Val | 390 | | | | | 395 | | | | | 400 |
| | | | | | Leu 405 | | | | | 410 | | | | | 415 | |
| | | | | 420 | Pro | | | | 425 | | | | | 430 | | |
| 20 | | | 435 | | Ala | | | 440 | | | | | 445 | | | |
| | | 450 | | | Val | | 455 | | | | | 460 | | | | |
| 25 | 465 | | | | Phe | 470 | | | | | 475 | | | | | 480 |
| | | | | | Gly 485 | | | | | 490 | | | | | 495 | |
| | - | | | 500 | Leu | | | | 505 | | | | | 510 | | |
| 30 | | | 515 | | Leu | | | 520 | | | | | 525 | | | |
| | | 530 | | | Glu | | 535 | | | | | 540 | | | | |
| 35 | 545 | | | | Gly | 550 | | | | | 555 | | | | | 560 |
| | | | | | Pro 565 | | | | | 570 | | | | | 575 | |
| •• | | | | 580 | Leu | | | | 585 | | | | | 590 | | |
| 40 | | | 595 | | Asn | | | 600 | | | | | 605 | | | |
| | | 610 | _ | _ | Pro | | 615 | | | | | 620 | | | | |
| 45 | 625 | | | | Thr | 630 | | | | | 635 | | | | | 640 |
| | | | | | Gly 645 | | | | | 650 | | | | | 655 | |
| 5 0 | | | | 660 | | | | | 665 | | | | | 670 | | |
| 50 | | | 675 | | Gln | | _ | 680 | | _ | | | 685 | | | |
| | | 690 | | | Val | | 695 | | | | | 700 | | | | |
| 55 | 705 | | | | | 710 | | | | | 715 | | | | | 720 |
| | | | | | His 725 | | | | | 730 | | | | | 735 | |
| CO | - | | | 740 | | | | | 745 | | | | | 750 | | |
| 60 | - | | 755 | | Ile | | | 760 | _ | | | | 765 | | | |
| | | 770 | | | Ser | | 775 | | | | | 780 | | | | |
| 65 | 785 | | | | Leu | 790 | | | | | 795 | | | | | 800 |
| | | | | | Asn 805 | | | | | 810 | | | | | 815 | |
| | | | | 820 | | | | | 825 | | | | | 830 | | |
| 70 | T).t | Ile | Ser 835 | _ | Asn | Lys | Ile | Val 840 | | Суѕ | Asp | Glu | Ala 845 | | Val | Leu |
| | Asn | Asn 850 | | Gly | Asn | Arg | Thr 855 | | Arg | Leu | His | Asn 860 | | Thr | Arg | Asn |
| 75 | Val 865 | Il∈ | | Asn | Cle | Arg 870 | | Gly | Ser | Thr | Leu 875 | | Asn | Ser | Tyr | 880 G17 |

```
Ile Tyr Asn Arg Asn Lys Ile Ser Asn Asn His Ile Gly Val Arg Leu
885 890 895
       Leu Asn Asn Ser Cys Phe Tyr Phe Asp Asn Ala Pro Val Ile Asn Glu
900 905 910
       Glu Asp Lys Gln Thr Phe Ile Ser Asn Arg Thr Trp Gln Leu Tyr Ser
915 920 925
 5
       Ser Asn Gly Thr Phe Pro Leu Asn Phe His Tyr Asn Ser Leu Gln Gly 930 935 940
       Gly Asp Thr Asp Thr Trp Ile Tyr Asn Asp Thr Tyr Thr Asn Arg Tyr 945 950 955 960
10
       Ile Asp Val. Ser Asn Asn His Trp Gly Asn Asn Asp Leu Phe Asp Pro
965 970 975
       Asn Gln Val Phe Asn Thr Pro Asp Leu Phe Ile Trp Ile Pro Phe Trp
980 985 990
       Asp Gly Leu Pro Asn Gly Arg Ser Gly Asn Ser Ser Ala Glu Ala Val
995 1000 1005
Glu Phe Gln Thr Ala Leu Asp Cys Ile Gly Asn Ser Asp Tyr Leu Ser
1010 1015 1020
15
       Ala Lys Val Ala Leu Lys Met Net Val Glu Thr Tyr Pro Glu Ser Asp
1025 1030 1035
20
       Phe Ala Ile Ala Ala Leu Lys Glu Leu Phe Arg Ile Glu Lys Met Ser
1045 1050 1055
       Gly Asn Asp Tyr Glu Gly Leu Lys Asp Tyr Phe Arg Ser Asn Pro Thr
1060 1065 1070
       Ile Ile Ser Ser Gln Asn Leu Phe Pro Thr Ala Asp Phe Leu Ser Ala
1075 1080 1085
25
       Arg Cys Asp Ile Val Cys Glu Asn Tyr Gln Ser Ala Ile Asp Trp Tyr
1090 1095 1100
       Glu Asn Arg Leu Asn Ser Glu Ile Ser Tyr Gln Asp Ser Val Phe Ala
1105 1110 1115
30
       Val Ile Asp Leu Gly Asp Ile Tyr Trp Asn Net Gln Leu Asp Ser Leu 1125 1130 1135
       Arg Gl; Thr Gly Ile Asp Leu Asn Ile Leu Ser Cys Glu Gln Arg Lys
1140 1145 1150
35
       Ser Leu Glu Ser His Gln Asn Val Lys Asn Tyr Leu Leu Ser Thr Leu
1155 1160 1165
       Pro Glu Ser Thr Gly Thr Leu Leu Pro Pro Leu Glu Cys Asn Lys Ser
1170 1175 1180
       Ser Leu Asp Lys Ser Lys Ile Ile Ser Ile Ser Pro Asn Pro Ala Lys
1185 1190 1195
40
       Ala Val Val Thr Ile Ile Tyr Tyr Thr Asp Asn Pro Ser Cys Ser Val
1205 1210 1215
       Ile Lys Ile Tyr Gly Ile Asn Gly Ala Ser Ala Asp Ile Thr Gly Leu
1220 1235 1230
       Pro Lys His Leu Ser Glu Gly Tyr Tyr Ser Ile Gln Phe Asn Thr Ser
1235 1240 1245
45
       Asn Phe Asp Pro Gly Phe Tyr Leu Val Thr Leu Asn Val Asp Gln Lys
1250 1260
                                1255
       Ile Ile Asp Thr Glu Lys Leu Arg Ile Lys
50
                               1270
       1265
        (2) INFORMATION FOR SEQ ID NO:467
              (i) SEQUENCE CHARACTERISTICS:
55
                   (A) LENGTH: 925 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
60
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
65
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...925
70
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:467
       Met Ala Ile Met Met Lys Ser Ile Val Phe Arg Ala Phe Leu Thr Ile
                                                  10
       Leu Leu Ser Trp Ala Ala Ile Thr Asn Pro Thr Ala Gln Glu Ile Ser
75
                                             25
```

| | G17. | Het | Aen 35 | Ala | Ser | Сув | Leu | Ala 40 | Ala | Pro | Ala | Gln | Pro 45 | Asp | Thr | lle |
|------------|------------|-----|------------|------------|------------|------------|-----|------------|------------|------------|------------|-------------|------------|------------|------------|------------|
| | | 50 | Glu | | | | 55 | G17. | | | | 60 | | | | |
| 5 | 65 | | Alā | | | 70 | | | | | 75 | | | | | 80 |
| | | | Vál | | 85 | | | | | 90 | | | | | 95 | |
| 10 | | | GlŢ | 100 | | | | | 105 | | | | | 110 | | |
| | | | Ala 115 | | | | | 120 | | | | | 125 | | | |
| 15 | | 130 | Glu | | | | 135 | | | | | 140 | | | | |
| 15 | 145 | | Phe Val | | | 150 | | | | | 155 | | | | | 160 |
| | | | Ala | | 165 | | | | | 170 | | | | | 175 | |
| 20 | | | Leu | 180 | | | | | 185 | | | | | 190 | | |
| | | _ | 195 Thr | | | | | 200 | | | | | 205 | | | |
| 25 | | 210 | Trp | | | | 215 | | | | | 220 | | | | |
| 20 | 225 | | Leu | | | 230 | | | | | 235 | | | | | 240 |
| | | | Gln | | 245 | | | | | 250 | | | | | 255 | |
| 30 | | | Leu | 260 | | | | | 265 | | | | | 270 | | |
| | | | 275 Asp | | | | | 280 | | | | | 285 | | | |
| 35 | | 290 | - | | | | 295 | | | | | 300 | | | | |
| | 305 | | G1; | | | 310 | | | | | 315 | | | | | 320 |
| | | | Tyr | | 325 | | | | | 330 | | | | Asn | 335 | |
| 40 | | | Ser | 340 | | | | Tyr | 345 | | | | Gly | 350 Pro | | |
| | Pro | | 355 Asn | туг | Leu | Ile | | | Lγε | Val | Glu | 380 G17. | 365 Ala | | Arg | Val |
| 4 5 | | | Trp | Val | Ser | Thr 390 | | | Ala | Asn | Trp 395 | | Ala | Glu | His | Tyr 400 |
| | 385 Ala | Val | Het | Ala | Ser 405 | | | G1; | Thr | Ala 410 | Val | Gly | Asp | Phe | Val 415 | |
| 50 | Leu | Phe | Glu | Glu 420 | Thr | Met | Thr | Ala | Lys 425 | Pro | | G1; | Als | Trp 430 | Tyr | Glu |
| | Arg | Thr | Il∈ 435 | | | Pro | Glu | G1; 440 | Thr | | Tyr | Ile | Ala 445 | | Arg | His |
| | | 450 | C).s | | | | 455 | | | | | 460 | | | | |
| 55 | 465 | | Thr | | | 470 | | | | | 475 | | | | | 480 |
| | | | Ile | | 485 | | | | | 490 | | | | | 495 | |
| 60 | | | Glu | 500 | | | | | 505 | | | | | 510 | | |
| | | | Asn 515 | | | | | 520 | | | | | 525 | | | |
| o.= | | 530 | | | | | 535 | | | | | 540 | | | | |
| 65 | 545 | , | . Glu | | | 550 | | | | | 555 | | | | | 560 |
| | | | Gln | | 565 | | | | | 570 | | | | | 575 | |
| 70 | | | Leu | 580 | | | | | 585 | | | | | 590 | | |
| | | | 11e 595 | | | | | 600 | 1 | | | | 605 | | | |
| 75 | _ | 610 | | _ | | | 615 | , | | | | 620 | | | | |
| 75 | T.7.1 | rec | ı Pro | Het | 116 | GT. | val | теп | Inr | Pro | ASP | ASD | тyr | ren | vāl | ınr |

630

```
Pro Arg Leu Glu Gly Ala Lys Leu Val Lys Tyr Trp Val Ser Ala Gln
645 650 655
       Asp Ala Val Tyr Ser Ala Glu His Tyr Ala Val Met Val Ser Thr Thr 660 665 670
       Gly Thr Ala Val Glu Asp Phe Val Leu Phe Glu Glu Thr Het Thr 675 680 685

Ala Lys Ala Asn Gly Ala Trp Tyr Glu Arg Thr Ile Thr Leu Pro Ala 690 695
       Gly Thr Lys Tyr Ile Ala Trp Arg His Tyr Asp Cys Thr Asp Met Phe 705 710 715 720
10
       Phe Leu Leu Asp Asp Ile Thr Val Tyr Arg Ser Thr Glu Thr Val 725 \phantom{\bigg|} 736 \phantom{\bigg|} 735
       Pro Glu Pro Val Thr Asp Phe Val Val Ser Leu Ile Glu Asn Asn Lys
740 745 750
15
       Gly Arg Leu Lys Trp Asn Tyr Pro Asn Gly Tyr Glu Pro Asp Lys Thr
755 760 765
       Asp Asp Lys Lys Pro Leu Gln Leu Thr Gly Tyr Asn Ile Tyr Ala Asn 770 780
20
       Gly Ser Leu Leu Val His Ile Gln Asp Pro Thr Val Leu Glu Tyr Ile
785 790 795 800
       Asp Glu Thr Tyr Ser Ser Arg Asp Gly Gln Val Glu Met Glu Tyr Cys
805 810 815
       Val Thr Ala Val Tyr Asn Asp Asn Ile Glu Ser Gln Ser Val Cys Asp
820 825 830
25
       Lys Leu Asn Tyr Thr Ile Thr Ser Leu Asp Asn Ile Gln Ser Asp Thr 835 840 845
       Ser Leu Lys Ile Tyr Pro Asn Pro Ala Ser Tyr Val Val Arg Ile Glu
850 855 860
30
       Gly Leu Ser Arg Ser Lys Ser Thr Ile Glu Leu Tyr Asn Ala Leu Gly
865 870 870 875
       The Cys lie Leu Arg Glu Glu Thr His Ser Glu Lys Thr Glu Ile Asp
885 890 895
       Val Ser Arg Leu Asn Asp Gly Val Tyr Leu Ile Lys Val Val Gly Gly 900 905 910
35
       Asn Lys Thr Thr Glu Lys Val Glu Ile Lys Arg Pro
       (2) INFORMATION FOR SEQ ID NO:468
40
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 922 amino acids
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
45
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
50
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
55
                   (B) LOCATION 1...922
            (xi) SEQUENCE DESCRIPTION: SEQ ID 110:468
       Het Met Lys Ser Ile Val Phe Arg Ala Phe Leu Thr Ile Leu Leu Ser
60
       Trp Ala Ala Ile Thr Asn Pro Thr Ala Gln Glu Ile Ser Gly Het Asn 20 25 30
       Ala Ser Cys Leu Ala Ala Pro Ala Gln Pro Asp Thr Ile Leu Tyr Glu
35 40 45
65
       Ser Phe Glu Asn Gly Pro Val Pro Asn Gly Trp Leu Glu Ile Asp Ala
50, 55 60
       Asp Ala Asp Gly Ala Thr Trp Gly Ser Pro Ser Gly Ser Phe Ser Val
65 70 75 80
       Pro Tyr Gly His Ash Gly Leu Cys Thr Tyr Ser His Ile Arg Ser Gly 85 90 95
70
       Ile Ser Thr Ala Gly Asn Tyr Leu Ile Thr Pro Asn Ile Glu Gly Ala
       Lys Arg Val Lys Tyr Trp Val Cys Asn Gln Tyr Ser Thr Asn Pro Glu
115 120 125
75
       His Tyr Ala Val Met Val Ser Thr Thr Gl; Thr Ala Ile Glu Asp Phe
```

| | | | | | | | | | | | | 1.10 | | | | |
|------------|------------|------------|--------------|------------|--------------|------------|------------|------------|------------|--------------|------------|------------|------------|------------|------------|--------------|
| | Val | 130 Leu | Leu | Phe | Asp | Asp | 135 Ser | Ile | Thr | Gly | Lys | 140 Pro | Thr | Pro | Leu | Val |
| _ | 1.45 | | Arg | | Ile | 150 | | | | Glu | 155 | | | | | 160 |
| 5 | Trp | Arg | His | Tyr 180 | 165 Lys | Vāl | Thr | Asp | Ser 185 | 170 His | Thr | Glu | ₽h∈ | Leu 190 | _ | Leu |
| | Asp | Asp | Val 195 | Thr | Val | Tyr | Arg | Ser 200 | Ile | Glu | Gly | Pro | Glu 205 | Pro | Ala | Thr |
| 10 | Asp | Phe 210 | Thr | Val | lle | Asn | 11e 215 | | Gln | Asn | Val | Gly 220 | Aṛg | Leu | Thr | Trp |
| | Asn 225 | Tyr | Pro | Glu | Asp | Туг 230 | Gln | Pro | Glu | Gly | Lys 235 | | Asn | Glu | Glu | Leu 240 |
| 15 | Gln | Leu | Ser | Glу | Tyr 245 | Asn | Ile | туг | Ala | Aen 250 | Gl; | Thr | Leu | Leu | Ala 255 | Gl n |
| 10 | lle | Lys | Asp | Val 260 | Ser | Ile | Leu | Glu | Tyr 265 | Val | Asp | Ser | Thr | Tyr 270 | Ser | Leu |
| | Arg | Asp | Asn 275 | Pro | Leu | Gln | Val | Glu 280 | Туг | Cγs | Val | Thr | Ala 285 | | Туг | Asp |
| 20 | Glu | Ser 290 | Ile | Glu | Ser | Ser | Thr 295 | | СЛЕ | Gly | Thr | Leu 300 | | Tjr | Ala | Thr |
| | Asp 305 | Ala | Ile | Leu | 4.7.L | Glu 310 | | Ph∈ | Glu | Asn | Gly 315 | | Val | Pro | Asn | G1; 320 |
| 25 | Trp | Fen | Val | Ile | Агр 325 | Ala | Asp | Gly | Asp | G1; 330 | Phe | Ser | Trp | G1?. | His 335 | туг |
| 23 | Leu | Asn | Ala | Туг 340 | Asp | Ala | Phe | Pro | Gly 345 | | Asn | Gly | G17. | His 350 | Сув | Ser |
| | Leu | Ser | Ala 355 | Ser | Tyr | Val | Pro | Gly 360 | | Glγ | Fro | Vāl | Thr 365 | | Asp | Aεn |
| 30 | | 370 | lle | | | | 375 | Glu | | | | 380 | | | | |
| | 385 | | Thr | | | 390 | | | | | 395 | | | | | 400 |
| 35 | Ala | Ser | Thr | | 405 | | | | | 410 | | | | | 415 | |
| | | | Het | 420 | | | | | 425 | | | | | 430 | | |
| | | | Pro 435 | Glu | Glу | | | 440 | | | | | 445 | | | |
| 4 0 | | 450 | Ile | | | | 455 | | | | | 460 | | | | |
| | 465 | Ala | Ser | | | 470 | | | | | 475 | | | | | 480 |
| 45 | Glu | Asn | Asn | Lys | Gly 485 | | Leu | Lys | Trp | Asn 490 | Tyr | Pro | Asn | Gl; | Туг 495 | Glu |
| | | | Lys | 500 | ı | | | | 505 | | | | | 519 | | |
| | | | Ala 515 | | | | | 520 | | | | | 525 | | | |
| 50 | | 530 | Tyr | | | | 535 | | | | | 540 |) | | | |
| | 545 | Glu | Tyr | | | 550 | | | | | 555 | | | | | 569 |
| 55 | Ser | Val | C).E | | 565 | | | | | 570 | • | | | | 5/5 | |
| | ТУT | Glu | ı Gly | Phe 580 | | Ala | Gl y | Ser | 11e 585 | | Glu | G13 | Trp | Leu 590 | Leu | Ile |
| | Asp | Ala | Asp 595 | | Asp | Asn | Val | Asn 600 | | Asp | Tr | туг | Fro 605 | Trp | Thr | Het |
| 60 | ጥን r | 613 | His | Asp | Ser | Glu | Lys 615 | | Ile | Ala | Ser | Pro 620 | Ser | Tyr | Leu | Pro |
| | 625 | : Ile | e G1; | | | 630 | Pro | Asp | | | 535 | , | | | | 640 |
| 65 | Glu | . G13 | Ala | Lys | Leu 645 | Val | Lys | Tyr | Tr | Val 650 | Ser | Ala | a Gln | Asp | Ala 655 | Val |
| | L7.1 | : Sei | . Ala | Glu 660 | ı His | | Ala | Val | Het 665 | : Val | | Thi | Thr | 670 | Thr | Ala |
| | ۷a] | l Glu | a Asp 675 | Phe | ≥ Val | Leu | Lei | Phe 680 | Glu | | ולד ו | : Het | Thr 685 | : Ala | Lys | Ala |
| 70 | Ası | o Gly | , Ala | Trp | о Туг | Glu | Arc 695 | Thr | | e Thi | : Let | Pro 700 | o Ala | | Thr | Lys |
| | Ty: | r Ile | ≜ Ala | Trp | Arç | His 710 | Туг | | су: | Thi | Asp 719 | Me1 | | Phs | . Ter | i Leu 720 |
| <i>7</i> 5 | Lei | n Vel | o Asp | ıl. | e Thi 725 | . Val | | r Arç | g Se | r Thi 730 | : Glu | | r Val | l Pro | Glu 735 | ı Pro |
| _ | | | | | | | | | | | | | | | | |

```
Val Thr Asp Phe Val Val Ser Leu Ile Glu Asn Asn Lys Gly Arg Leu 740 745 750
        Lys Trp Asn Tyr Pro Asn Gly Tyr Glu Pro Asp Lys Thr Asp Asp Lys
755 760 765
       Lys Pro Leu Gln Leu Thr Gly Tyr Asn Ile Tyr Ala Asn Gly Ser Leu
770 780
       Leu Val His Ile Gln Asp Pro Thr Val Leu Glu Tyr Ile Asp Glu Thr 785 790 795 800
       Tyr Ser Ser Arg Asp Gly Gln Val Glu Met Glu Tyr Cys Val Thr Ala
805 810 815
10
       Val Tyr Acn Asp Asn Ile Glu Ser Gln Ser Val Cys Asp Lys Leu Asn
820 825 830
       Tyr Thr Ile Thr Ser Leu Asp Asn Ile Gln Ser Asp Thr Ser Leu Lys
835 840 845
15
       Ile Tyr Pro Asn Pro Ala Ser Tyr Val Val Arg Ile Glu Gly Leu Ser
850 855 860
       Arg Ser Lys Ser Thr Ile Glu Leu Tyr Asn Ala Leu Gly Ile Cys Ile
865 870 875 880
                             870
        Leu Arg Glu Glu Thr His Ser Glu Lys Thr Glu Ile Asp Val Ser Arg
885 890 895
20
       Leu Asn Asp Gly Val Tyr Leu Ile Lys Val Val Gly Gly Asn Lys Thr 900 905 910
        Thr Thr Glu Lys Val Glu Ile Lys Arg Pro
25
        (2) INFORMATION FOR SEQ ID NO:469
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 921 amino acids
30
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
35
           (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
40
             (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...921
             (mi) SEQUENCE DESCRIPTION: SEQ ID NO:469
45
       Het L;s Ser Ile Val Phe Arg Ala Phe Leu Thr Ile Leu Leu Ser Trp 1 \phantom{\bigg|} 5 \phantom{\bigg|} 10 \phantom{\bigg|} 15
       Ala Ala Ile Thr Asn Pro Thr Ala Gln Glu Ile Ser Gly Met Asn Ala
20 25 30
50
       Ser Cys Leu Ala Ala Pro Ala Gln Pro Asp Thr lle Leu Tyr Glu Ser 35 40 45
       Phe Glu Asn Gly Pro Val Pro Asn Gly Trp Leu Glu Ile Asp Ala Asp
50 55 60
       Ala Asp Gly Ala Thr Trp Gly Ser Pro Ser Gly Ser Phe Ser Val Pro 65 70 75 80
55
       Tyr Gly His Asn Gly Leu Cys Thr Tyr Ser His Ile Arg Ser Gly Ile
85 90 95
Ser Thr Ala Gly Asn Tyr Leu Ile Thr Pro Asn Ile Glu Gly Ala Lys
100 105
       Arg Val Lys Tyr Trp Val Cys Asn Gln Tyr Ser Thr Asn Pro Glu His
115 120 125

Tyr Ala Val Met Val Ser Thr Thr Gly Thr Ala Ile Glu Asp Phe Val
130 135
60
       Leu Leu Phe Asp Asp Ser Ile Thr Gly Lys Pro Thr Pro Leu Val Trp
145 150 150 160
65
       Arg Arg Arg Ile Val Asp Leu Pro Glu Gly Thr Lys Tyr Ile Ala Trp
165 170 175
       Arg His Tyr Lys Val Thr Asp Ser His Thr Glu Phe Leu Lys Leu Asp 180 185 190
70
       Asp Val Thr Val Tyr Arg Ser Ile Glu Gly Pro Glu Pro Ala Thr Asp
195 200 205
       Phe Thr Val Ile Asn Ile Gly Gln Asn Val Gly Arg Leu Thr Trp Asn 210 215 220
        Tyr Pro Glu Asp Tyr Gln Pro Glu Gly Lys Gly Asn Glu Glu Leu Gln
75
```

| | Leu | Ser | Gly | Tyr | Asn 245 | Ile | T).r | Ala | Asn | G1 y 250 | Thr | Leu | Leu | Ala | Gln 255 | Ile |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|--------------------|------------|------------|
| | L;:s | Asp | Val | Ser 260 | Ile | Leu | Glu | Туr | Val 265 | | Ser | Thr | ТУr | Ser 270 | | Arg |
| 5 | Asp | Asn | Pro 275 | Leu | Gln | Vāl | Glu | Tyr 280 | | Val | Thr | Ala | Val 285 | | Asp | Glu |
| | Ser | 11e 290 | Glu | Ser | Ser | Thr | Val 295 | Сув | Glу | Thr | Leu | His 300 | | Ala | Thr | Asp |
| 10 | Ala 305 | Ile | Leu | T;·r | Glu | Asn 310 | | Glu | Asn | G1? | Pro 315 | Vāl | Pro | Asn | Gly | Trp 320 |
| 10 | Leu | Val | Ile | Asp | Ala 325 | Asp | Gl7. | Asp | $G1\gamma$ | Phe 330 | | Trp | Gly | His | Tyr 335 | Leu |
| | Asn | Ala | Tyr | Asp 340 | | Phe | Pro | Gly | His 345 | | Glу | Gly | His | Су <i>в</i> 350 | Ser | Leu |
| 15 | Ser | Ala | Ser 355 | Тут | Val | Pro | Glγ | 11e 360 | Glγ | Pro | Val | Thr | Pro 365 | Asp | Asn | Туг |
| | Leu | 11e 370 | Thr | Pro | Lys | Vāl | Glu 375 | Gl y | Ala | Lys | Arg | Val 380 | Lys | T'r | Trp | Val |
| 20 | Ser 385 | Thr | Gln | Asp | Ala | Asn 390 | Trp | Ala | Ala | Glu | His 395 | T;r | Alā | Val | Het | Ala 400 |
| | Ser | Thr | Thr | GlŢ | Thr 405 | | Val | Gly | Asp | Phe | | Ile | Leu | Phe | Glu 415 | Glu |
| | Thr | Met | Thr | Ala 420 | | Pro | Thr | Gly | Ala 425 | | Туг | Glu | Arg | Thr 430 | Ile | Asn |
| 25 | Leu | Pro | Glu 435 | Glγ | Thr | Lys | Tyr | 11e 440 | | Trp | Arg | His | Tyr 445 | Asn | C?.e | Thr |
| | Asp | Ile 450 | Tyr | Phe | Leu | Γλ.ε | Leu 455 | Asp | Asp | Ile | Thr | Val 460 | Phe | Gly | Thr | Pro |
| 30 | 465 | Ser | | | | 470 | | Thr | | | 475 | | | | | 480 |
| | Asn | Asn | Lys | Gly | Arg 485 | Leu | Lys | Trp | Asn | Tyr 490 | Pro | Asn | Gly | Tyr | Glu 495 | Pro |
| | Asp | Lys | Thr | Asp 500 | Asp | L;•s | Asp | Pro | Leu 505 | Gln | Leu | Ala | Gly | T;r 510 | Asn | Ile |
| 35 | Туr | Ala | Asn 515 | Gly | Ser | Leu | Leu | Val 520 | His | Ile | Gln | Asp | Pro 525 | Thr | Val | Leu |
| | | 530 | | | | | 535 | | | | | 540 | | | | |
| 40 | 545 | Tyr | Cys | | | 550 | | Tyr | | | 555 | | | | | 560 |
| | | | | | 565 | | | Asp | | 570 | | | | | 5/5 | |
| | | | | 580 | | | | Ile | 585 | | | | | 590 | | |
| 45 | | | 595 | | | | | Trp 600 | | | | | 605 | | | |
| | | 610 | | | | | 615 | | | | | 620 | | | | |
| 50 | 625 | | | | | 630 | | | | | 635 | | | | | Glu 640 |
| | | | | | 645 | | | Trp | | 650 | | | | | 655 | |
| | | | | 660 | | | | | 665 | | | | | 670 | | Vāl |
| 55 | | | 675 | | | | | 680 | | | | | 685 | | | Asn |
| | | 690 | | | | | 695 | | | | | 700 | | | | Tyr |
| 60 | 705 | | | | | 710 | | | | | 715 | | | | | Leu 720 |
| | | | | | 725 | | | | | 730 | 1 | | | | 735 | |
| | | | | 740 | | | | | 745 | | | | | 750 | | Lys |
| 65 | | | 755 | | | | | 760 | | | | | 765 | | | r7.2 |
| | | 770 | i | | | | 775 | , | | | | 780 | | | | Leu |
| 70 | 785 | ı | | | | 790 | | | | | 795 | | | | | Tyr 800 |
| | | | | | 805 | | | | | 810 |) | | | | 815 | |
| | _ | | _ | 820 | 1 | | | | 825 | | - | | | 830 | | Tyr |
| <i>7</i> 5 | Thr | Ile | Thr | Ser | Leu | Asp | Asr | ılle | Gl.n | Ser | Asp | Thr | Ser | Leu | Lys | 11- |

```
835
                                                                  845
                                         840
       Tyr Pro Asn Pro Ala Ser Tyr Val Val Arg Ile Glu Gly Leu Ser Arg
850 855 860
       Ser Lys Ser Thr Ile Glu Leu Tyr Asn Ala Leu Gly Ile Cys Ile Leu
                                                     875
       865
                              870
       Arg Glu Glu Thr His Ser Glu Lys Thr Glu Ile Asp Val Ser Arg Leu
885 890 895
Asn Asp Gly Val Tyr Leu Ile Lys Val Val Gly Gly Asn Lys Thr Thr
900 905 910
10
       Thr Glu Lys Val Glu Ile Lys Arg Pro
       (2) INFORMATION FOR SEQ ID NO:470
15
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 593 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
20
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
25
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...593
30
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:470
       Het Asn Ser Ile Met Lys Tyr Gln Leu Tyr Thr Ala Val Ile Het Ala
35
       Leu Ser Val Ser Ser Val Cys Gly Gln Thr Pro Arg Asn Thr Glu Thr 25 30
       Lys Arg Pro Asp Thr Leu Arg Arg Glu Leu Thr Ile Val Asn Asp Gln 35 40 45
       Thr Val Glu Het Glu His Ala Asp Pro Leu Pro Ala Ala Tyr Lys Ala
50 55 60
40
       Ile Glu Pro Arg Leu Lys Pro Phe Arg Pro Glu Tyr Asn Lys Arg Thr 65 70 75 80
       Phe Gly Phe Val Pro Glu Val Ser Ser Ser Gly Arg Asn Asn Leu Pro 85 \phantom{\bigg|}90\phantom{\bigg|}
45
       Asn Ile Leu Pro Thr Glu Gly His Met Lys His Arg Gly Tyr Leu Asn 100 \hspace{1cm} 105 \hspace{1cm} 110
       Ile Gly Ile Gly His Thr Leu Asn Gln Arg Het Asp Ala Gly Tyr Arg
115 120 125
       Leu Ile Asp Ala Glu Gln Glu Arg Leu Asn Leu Phe Leu Ser Tyr Arg
130 135
50
       Gly Net Lys Ser Ala Phe Asn Thr Gly Asp Phe Asp Gly Asp Arg Lys
145 150 155 160
       Asp Arg Arg Met Met Ala Gly Val Asp Tyr Glu Gln Arg Arg Pro Ser
165 170
       Phe Val Leu Ala Thr Gly Leu Tyr Tyr Ser Asn His Tyr Phe Asn Asn 180 190
       Tyr Gly Arg Gly Ala Thr Thr Asn Val Gly Ser Ile Pro Gln Leu Ser
195 200 205
       Thr Pro Val Thr Pro Gln Net Asp Asn Gly Thr His Asn Val Arg Val 210 215 220
       Tyr Leu Gly Ala Lys Asn Asp Val Ile Asp Ala Arg Ile Asp Tyr Arg
225 230 230 235
       Phe Phe Arg Ser Ile Pro Tyr Leu Gly Thr Asp Pro Met Lys Ala Leu
245 250 355
65
       Thr Glu His Thr Pro Glu Leu Asn Val Thr Net Ser Asn Glu Leu Ser 260 265 270
       Asp Asp Ile Lys Leu Gly Val Glu Val Arg Thr Gly Gly Leu Fhe Phe 275 280 285
       Ala Lys Asn Ser Glu Het Ile Gln Thr Gly Val Leu Ser Glu Thr Asp 290 295 300

Arg Asn Leu Tyr Tyr Val Glu Gly Ala Pro Thr Ile Gly Phe Val Gl; 305 310 320
70
       Asp Ser Asp Asn Het Gln Trp Asn Ile Gln Ala Gly Val Gl; Ile Ser
                                                 330
       Ser His Phe Gly Ala Lys Gly Arg Leu Phe Phe Trp Pro Lys Leu Asp
```

```
350
                                                    345
                        340
        Ala Ser Leu Ser Ile Phe Pro Ser Trp Arg Val Tyr Ala Lys Ala Phe 355 360 365

Gly Gly Val Ile Arg Asn Gly Leu Ala Asp Val Het Gln Glu Het 370 380
        Pro Tyr Leu Met Pro Asn Thr Ile Val Leu Pro Ser Arg Asn Ala Leu
385 390 395 400
        Thr Ala Gln Leu Gly Val Lys Gly Asn Ile Ala Asp Val Val Arg Het
405 415
        Glu Val Tyr Gly Asp Phe Ser Lys Leu Thr Gly Val Pro Phe Tyr Thr 420 435
10
        Pro Thr Leu Pro Leu Tyr Asn Pro Ser Asp Leu Tyr Gln Tyr Asn Val
435 440
        Ser Phe Leu Pro Ile Tyr Ala Asp Gly Ser Arg Trp Arg Ala Gly Gly
455 460
15
        Lys Leu Glu Tyr Ser Tyr Arg Asp Met Leu Arg Phe Leu Val Asp Ala
465 470 475 480
        Ser Tyr Gly Lys Trp Asn Leu Asp Gly Gly Leu Val Ala Ser Het Gln
485 495

        Pro
        Asp
        Leu
        Leu
        Lys
        Ala
        Glu
        Val
        Gly
        Val
        His
        Pro
        Ile
        Ala
        Pro

        Leu
        Asp
        Val
        Arg
        Leu
        Arg
        Tyr
        Thr
        Gln
        Leu
        Asp
        Arg
        Tyr
        Arg
        Tyr

        Leu
        Asp
        Val
        Arg
        Tyr
        Arg
        Tyr
        Arg
        Tyr

20
        Ser Phe Gly Ser Ala Gly Ser Glu Ala Leu Gly Ile Gly Asn Val His
530 535 540
25
        Leu Leu Ser Ala Asp Val Ser Tyr Lys Leu Lys Lys Asn Leu Ser Leu
545 550 555 560
        Tyr Leu Lys Ile Asp Asn Net Leu Ala Glu Thr Thr Glu Leu Ile Gly
565 575
        Tyr Tyr Pro Het Gln Pro Phe His Cys Phe Ala Gly Phe Ser Trp Thr
580 585
30
         (2) INFORMATION FOR SEQ ID NO:471
35
               (i) SEQUENCE CHARACTERISTICS:
                      (A) LENGTH: 589 amino acids
                      (B) TYPE: amino acid
40
                      (D) TOPOLOGY: linear
              (ii) MOLECULE TYPE: protein
             (iii) HYPOTHETICAL: YES
45
              (vi) ORIGINAL SOURCE:
                      (A) ORGANISH: Porphyromonas gingivalis
              (ix) FEATURE:
50
                      (A) HAME/KEY: misc feature
                      (B) LOCATION 1...5\overline{8}9
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 471
        Het Lys Tyr Gln Leu Tyr Thr Ala Val Ile Het Ala Leu Ser Val Ser 1 5 10 15
55
         Ser Val Cys Gly Gln Thr Pro Arg Asn Thr Glu Thr Lys Arg Pro Asp
20 25 30
         Thr Leu Arg Arg Glu Leu Thr Ile Val Asn Asp Gln Thr Val Glu Het
35 40 45
60
         Glu His Ala Asp Pro Leu Pro Ala Ala Tyr Lys Ala Ile Glu Pro Arg
50 55 60
        65
         His Thr Leu Asn Gin Arg Met Asp Ala Gly Tyr Arg Leu Ile Asp Ala
115 120 125
70
         Glu Gln Glu Arg Leu Asn Leu Phe Leu Ser Tyr Arg Gly Het Lys Ser
130 135 140
         Ala Phe Asn Thr Gly Asp Phe Asp Gly Asp Arg Lys Asp Arg Arg Net
145 150 155 160
75
         Met Ala Gly Val Asp Tyr Glu Gln Arg Arg Pro Ser Phe Val Leu Ala
```

```
165
                                                  170
       Thr Gly Leu Tyr Tyr Ser Asn His Tyr Phe Asn Asn Tyr Gly Arg Gly
180 185 190
       Ala Thr Thr Asn Val Gly Ser Ile Pro Gln Leu Ser Thr Pro Val Thr
 5
                                     200
               195
                                                               205
       Pro Gln Het Asp Asn Gly Thr His Asn Val Arg Val Tyr Leu Gly Ala 210 215 220
       Lys Asn Asp Val Ile Asp Ala Arg Ile Asp Tyr Arg Phe Phe Arg Ser 225 \phantom{\bigg|}230\phantom{\bigg|}230\phantom{\bigg|}235\phantom{\bigg|}235\phantom{\bigg|}
10
       Ile Pro Tyr Leu Gly Thr Asp Pro Net Lys Ala Leu Thr Glu His Thr 245 250 255
       Pro Glu Leu Asn Val Thr Met Ser Asn Glu Leu Ser Asp Asp Ile Lys
260 265 270
       Leu Gly Val Glu Val Arg Thr Gly Gly Leu Phe Phe Ala Lys Asn Ser 275 280 285
15
       Glu Het Ile Gln Thr Gly Val Leu Ser Glu Thr Asp Arg Asn Leu Tyr
290 295 300
       Tyr Val Glu Gly Ala Pro Thr Ile Gly Phe Val Gly Asp Ser Asp Asn 305 310 315 320
20
       Met Gln Trp Asn Ile Gln Ala Gly Val Gly Ile Ser Ser His Phe Gly 325 330 335
       Ala Lys Gly Arg Leu Phe Phe Trp Pro Lys Leu Asp Ala Ser Leu Ser 340 350
       Ile Phe Pro Ser Trp Arg Val Tyr Ala Lys Ala Phe Gly Gly Val Ile
355 360 365
25
       Arg Asn Gly Leu Ala Asp Val Met Gln Glu Glu Het Pro Tyr Leu Met 370 380
       Pro Asn Thr Ile Val Leu Fro Ser Arg Asn Ala Leu Thr Ala Gln Leu
385 390 395 400
30
       Gly Val Lys Gly Asn Ile Ala Asp Val Val Arg Het Glu Val Tyr Gly
405 415
       Asp Phe Ser Lys Leu Thr Gly Val Pro Phe Tyr Thr Pro Thr Leu Pro 420 425 430
       Leu Tyr Asn Pro Ser Asp Leu Tyr Gln Tyr Asn Val Ser Phe Leu Pro 435 440 445
35
       Ile Tyr Ala Asp Gly Ser Arg Trp Arg Ala Gly Gly Lys Leu Glu Tyr 450 460
       Ser Tyr Arg Asp Net Leu Arg Phe Leu Val Asp Ala Ser Tyr Gly Lys
465 470 475 480
40
       Trp Asn Leu Asp Gly Gly Leu Val Ala Ser Met Gln Pro Asp Leu Ile
485 490 495
                                               490
       Leu Lys Ala Glu Val Gly Val His Pro Ile Ala Pro Leu Asp Val Arg
500 505 510
       Leu Arg Tyr Thr Gln Leu Asn Gly Arg Tyr Arg Tyr Ser Phe Gly Ser 515 520 525

Ala Gly Ser Glu Ala Leu Gly Ile Gly Asn Val His Leu Leu Ser Ala 530 535
45
       Asp Val Ser Tyr Lys Leu Lys Lys Asn Leu Ser Leu Tyr Leu Lys Ile
                             550
                                            555
50
       Asp Asn Het Leu Ala Glu Thr Thr Glu Leu Ile Gly Tyr Tyr Pro Met 565 570 575
       Gln Pro Phe His Cys Phe Ala Gly Phe Ser Trp Thr Phe
55
        (2) INFORMATION FOR SEQ ID NO: 472
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 346 amino acids
                    (B) TYPE: amino acid
60
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
65
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
70
                   (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...346
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:472
75
       Net Net Glu Lys Cys Ile Phe Ala His Tyr Pro His Asn Leu Val Phe
```

```
Met Ile Arg Lys His Phe Gly Ile Ile Leu Gly Phe Leu Ser Leu Val
20 25 30
       Phe Ser Ala Gly Ala Gln Gln Glu Lys Gln Val Phe His Phe Leu Asn 35
 5
       Leu Pro Ala Thr Ala Gln Ala Leu Ala Ala Gly Gly Lys Ala Ile Thr 50 60
       The Val Asp Asp Asn Pro Gly Leu Ala Phe Glu Asn Pro Ala Leu Leu 65 70 75 80
       Gly Tyr Glu Ser Gly Gly Arg Ala Phe Leu Ser Tyr Leu Tyr Tyr Het 85 90 95
10
       Ser Gly Ser His Het Gly Asn Ala Cys Tyr Ala Ser Ser Val Gl; Glu
100 105 110
       Arg Gly Met Trp Gly Val Gly Met Arg Phe Leu Asn Tyr Gly Ser Het 115 120 125
15
       Gln Gly Tyr Asp Gln Asn Ala Ile Ala Thr Gly Ser Phe Ser Ala Ser
130 135 140
       Asp Ile Ala Val Gln Gly Phe Tyr Ser His Glu Leu Ser Asn His Phe 145 150 155 160

Arg Gly Gly Val Ser Leu Lys Ala Leu Tyr Ser Ser Ile Glu Thr Tyr 165 175 175
20
       Ser Ser Phe Gly Leu Gly Val Asp Val Gly Ile Ser Tyr Tyr Asp Asp 180 $190$
       Asp Lys Gly Tyr Ser Ala Ser Ala Leu Phe Lys Asn Val Gly Ala Gln
195 200 205
25
       Leu Lys Gly Tyr Asn Glu Glu Arg Glu Pro Leu Asp Trp Asp Phe Gln 210 215 220
       Leu Gly Phe Ser Arg Ser Phe Ile Asn Ala Pro Phe Arg Leu His Ile
225 230 235
       Thr Leu Phe Asn Leu Asn Pro His Tyr Phe Lys Arg Leu Val Pro Arg
245 250 255
30
       Asp Leu Ser Lys Met Gln Lys Phe Leu Arg His Phe Ser Ile Gly Ala 260 265 270
       Glu Phe Thr Pro Ser Glu Arg Phe Trp Val Gly Leu Gly Tyr Thr Pro
275 280 285
35
       Gln Ile Ala Gln Asp Phe Glu Val Glu Gly Gly Asn Lys Trp Gly Gly
290 295 300
       Leu Ser Ala Gly Val Gly Phe Thr Ser Gly Val Val Arg Val Gly Val 305 310 315 320
       Ser Ala Ala Thr Tyr His Pro Ala Ala Leu Ser Phe Het Cys Ser Val
325 330 335
40
        Gly Ile Arg Leu Asp Asp Lys Ser Ile Phe
                     340
45
        (2) INFORMATION FOR SEQ ID NO:473
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 345 amino acids
                    (B) TYPE: amino acid
50
                    (D) TOPOLOGY: linear
             (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
55
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISH: Porphyromonas gingivalis
             (ix) FEATURE:
60
                    (A) HAME/KEY: misc_feature
                    (B) LOCATION 1...345
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 473
65
        Het Glu Lys Cys Ile Phe Ala His Tyr Pro His Asn Leu Val Phe Het
                                                  10
        The Arg Lys His Phe Gly Ile Ile Leu Gly Phe Leu Ser Leu Val Phe 20 25 30
        Ser Ala Gly Ala Gln Gln Glu Lys Gln Val Phe His Phe Leu Asn Leu 35 40 45
 70
        Pro Ala Thr Ala Glm Ala Leu Ala Ala Gly Gly Lys Ala Ile Thr Ile
50 60
        Val Asp Asp Asn Pro Gly Leu Ala Phe Glu Asn Pro Ala Leu Leu Gly 65 70 75 80
 75
        Tyr Glu Ser Gly Gly Arg Ala Phe Leu Ser Tyr Leu Tyr Tyr Met Ser
```

90

```
Gly Ser His Met Gly Asn Ala Cys Tyr Ala Ser Ser Val Gly Glu Arg
100 105 110
       Gly Het Trp Gly Val Gly Met Arg Phe Leu Asn Tyr Gly Ser Met Gln
115 120 125
       Gly Tyr Asp Gln Asn Ala Ile Ala Thr Gly Ser Phe Ser Ala Ser Asp
130 135 140
       Ile Ala Val Gln Gly Phe Tyr Ser His Glu Leu Ser Asn His Phe Arg
145 150 160
       Gly Gly Val Ser Leu Lys Ala Leu Tyr Ser Ser Ile Glu Thr Tyr Ser 165 170 175

Ser Phe Gly Leu Gly Val Asp Val Gly Ile Ser Tyr Tyr Asp Asp Asp 180 185 190
10
       Lys Gly Tyr Ser Ala Ser Ala Leu Phe Lys Asn Val Gly Ala Gln Leu 195 200 205
15
       Lys Gly Tyr Ash Glu Glu Arg Glu Pro Leu Asp Trp Asp Phe Gln Leu 210 215 220
       Gly Phe Ser Arg Ser Phe Ile Asn Ala Pro Phe Arg Leu His Ile Thr
225 230 235 240
20
       Leu Phe Asn Leu Asn Pro His Tyr Phe Lys Arg Leu Val Pro Arg Asp 245 250 255
       Leu Ser Lys Net Gln Lys Phe Leu Arg His Fhe Ser Ile Gly Ala Glu
260 265 270
       Phe Thr Pro Ser Glu Arg Phe Trp Val Gly Leu Gly Tyr Thr Pro Gln 275 280 285
25
       Ile Ala Gin Asp Phe Glu Val Glu Gly Gly Asn Lys Trp Gly Gly Leu
290 295 300
       Ser Ala Gly Val Gly Phe Thr Ser Gly Val Val Arg Val Gly Val Ser 305 310 320
30
       Ala Ala Thr Tyr His Pro Ala Ala Leu Ser Phe Het Cys Ser Val Gly
       Ile Arg Leu Asp Asp Lys Ser Ile Phe
                     340
35
       (2) INFORNATION FOR SEQ ID NO: 474
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 330 amino acids
                   (B) TYPE: amino acid
40
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
45
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
50
                   (B) LOCATION 1...330
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:474
55
       Met Ile Arg Lys His Phe Gly Ile Ile Leu Gly Phe Leu Ser Leu Val
       Fhe Ser Ala Gly Ala Gln Gln Glu Lys Gln Val Phe His Phe Leu Asn 20 25 30
       Leu Pro Ala Thr Ala Gl<br/>n Ala Leu Ala Ala Gly Gly Lys Ala Ile Thr 35 \phantom{-}40\phantom{0} 45
60
       Ile Val Asp Asp Asn Pro Gly Leu Ala Phe Glu Asn Pro Ala Leu Leu 50 55 60
       Gly Tyr Glu Ser Gly Gly Arg Ala Phe Leu Ser Tyr Leu Tyr Tyr Het 65 70 75 80
65
       Ser Gly Ser His Met Gly Asn Ala Cys Tyr Ala Ser Ser Val Gly Glu 85 90 95
       Arg Gly Net Trp Gly Val Gly Met Arg Phe Leu Asn Tyr Gly Ser Net
100 105 110
       Gln Gly Tyr Asp Gln Asn Ala IIe Ala Thr Gly Ser Phe Ser Ala Ser 115 120 125

Asp IIe Ala Val Gln Gly Phe Tyr Ser His Glu Leu Ser Asn His Phe 130 135 140
70
       Arg Gly Gly Val Ser Leu Lys Ala Leu Tyr Ser Ser Ile Glu Thr Tyr
                            150
                                                 155
       Ser Ser Phe Gly Leu Gly Val Asp Val Gly Ile Ser Tyr Tyr Asp Asp
```

```
170
                        165
      Asp Lys Gly Tyr Ser Ala Ser Ala Leu Phe Lys Asn Val Gly Ala Gln
180 185 190
      Leu Lys Gly Tyr Asn Glu Glu Arg Glu Pro Leu Asp Trp Asp Phe Gln 195 200 205
 5
      Leu Gly Phe Ser Arg Ser Phe Ile Asn Ala Pro Phe Arg Leu His Ile
210 215 220
       Thr Leu Phe Asn Leu Asn Pro His Tyr Phe Lys Arg Leu Val Pro Arg
230 235 240
      Asp Leu Ser Lys Het Gln Lys Phe Leu Arg His Phe Ser Ile Gly Ala
245 250 255
10
       Glu Phe Thr Fro Ser Glu Arg Phe Trp Val Gly Leu Gly Tyr Thr Pro
260 265 270
       Gln Ile Ala Gln Asp Phe Glu Val Glu Gly Gly Asn Lys Trp Gly Gly 275 280 285
15
       Leu Ser Ala Gly Val Gly Phe Thr Ser Gly Val Val Arg Val Gly Val 290 295 300
       Ser Ala Ala Thr Tyr His Pro Ala Ala Leu Ser Phe Met Cys Ser Val
305 310 315 320
       Gly Ile Arg Leu Asp Asp Lys Ser Ile Phe
20
       (2) INFORMATION FOR SEQ ID NO:475
25
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 324 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
30
           (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
35
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) MAME/KEY: misc_feature
                   (B) LOCATION 1...324
40
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:475
       Net Cys Leu Glu Pro Ile Ile Ala Pro Ile Ser Ser Glu Leu Leu Glu 1 5 10 15
       Gln Glu Leu Thr Ala Asp Arg Phe Leu Arg Het Thr Asn Lys Ala Gly
20 25 30
45
       Asn Glu Ile Tyr Val Phe Thr Ala Glu Glu Ala Pro His Cys Met Lys
35 40
       Glu Val Gly Arg Leu Arg Glu Glu Ala Phe Arg His Tyr Gly Gly Gly 50 60
50
       Thr Gly Lys Ala Ile Asp Ile Asp Glu Phe Asp Thr Het Pro Gly Ser 65 70 75
       Tyr Lys Gln Leu Ile Val Trp Asp Pro Gln Asn Lys Ala Ile Leu Gly
85 90 95
       Gly Tyr Arg Phe Ile Tyr Gly Arg Asp Val Ala Phe Asp Thr Asp Gly 100 105
55
        Lys Pro Leu Leu Ala Thr Ala Glu Net Phe Arg Phe Ser Asp Ala Phe
115 120 125
        Leu His Asp Tyr Leu Pro Tyr Thr Val Glu Leu Gly Arg Ser Phe Val
130 135 140
 60
        Ser Leu Gln Tyr Gln Ser Thr Arg Met Gly Thr Lys Ala Ile Phe Val
145 150 155 160
        Leu Asp Asn Leu Trp Asp Gly Ile Gly Ala Leu Thr Val Val Asn Pro
165 176 175
        Glu Ala Leu Tyr Phe Tyr Gly Lys Val Thr Het Tyr Lys Asp Tyr Asp 180 185 190
 65
        Arg Arg Ala Arg Asn Leu Ile Leu Tyr Fhe Leu Arg Lys His Phe Ser
        Asp Pro Glu Gly Leu Val Lys Pro Ile His Pro Leu Pro Ile Glu Ile
210 215 220
 70
        Ser Ala Glu Asp Glu Ala Leu Phe Ser Ser Ser Asp Phe Asp Thr Asn
225 230 235 240
        Tyr Lys Thr Leu Asn Ile Glu Val Arg Lys Leu Gly Ile Asn Ile Pro
245 255
        Pro Leu Val Ser Ala Tyr Ile Ala Leu Ser Pro Glu Het Arg Val Phe
 75
```

```
Gly Thr Ala Val Asn Glu Ser Phe Gly Glu Val Glu Glu Thr Gly Ile
275 280 285
                                      280
       Phe Ile Ala Val Gly Lys Ile Leu Glu Glu Lys Lys Gln Arg His Ile
 5
                                 295
                                                      300
       Glu Ser Phe Ile Leu Ser Arg Asn Glu Lys Lys Gly Leu Asp Ser Ser
       305
       Asn Gly Arg Ser
10
       (2) INFORMATION FOR SEQ ID NO:476
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 547 amino acids
15
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
20
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
25
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...547
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:476
30
       Het Lys Thr Ile Val Arg Tyr Ser Arg Leu Pro Val Ala Leu Phe Phe
                                               10
       Cys Leu Leu Gly Ala Val His Leu Ser Val Glu Ala Gln Met Leu Asn
20 25 30
35
       Thr Pro Phe Glu Leu Ser Asp Gln Ile Val Leu Ser Pro Thr Glu Arg
                                     40
       Gin Tyr Arg Glu Ile Cys Val Gln Thr Lys Glu Lys Arg Gly Ala Asp
50 60
       Leu Phe Pro Leu Ser Asp Lys Leu Arg Asp Ser Ala Tyr Val Arg Phe 65 70 80
40
       Gly Ser Ala Tyr Gly Asp Ile Ala Gly Asp Tyr Leu Pro Tyr Asn Gly
85 90 95
       Acn Asn Tyr Ser Ser Leu Ser Leu Glu Ser Gly Gly Arg Ile Ser Val 100 \, 105 \, 110 \,
45
       Arg Asn Tyr Gly Thr Leu Gln Gly Ser Ala Ser Tyr Ser Arg Gly Met
115 120 125
       His Lys Arg Ile Gly Trp Asn Ala Leu Arg Asn Ala Glu Ala Tyr Tyr
130 135
       Pro Tyr Leu Val Ser Asp Ser Thr Gly Gly Asp Tyr His Phe Glu Asp
145 150 155 160
50
       Tyr Arg Leu Ala Gly Tyr Tyr Ser Phe Arg Ala Gly Arg Leu Pro Leu
165 170
       Gly Ile Gly Phe Ser Tyr Arg Gly Glu Val Ala Tyr Arg Leu Thr Asp
180 185 190
55
       Pro Arg Thr Thr Asn Thr Thr Gly Ala Leu Glu Leu Ser Cys Ala Thr
195 200 205
       Ser Leu Thr Leu Pro Arg Glu Asn Arg Leu Ser Leu Ser Ala Ala Tyr
210 215 220
       Leu Tyr His Arg Gln His Leu Thr Gln Tyr Asn Trp Arg Pro Gly Gln 225 230 230 240
60
       Gln Asp Lys Phe Phe Val Ser Tyr Gly Phe Gly Gln Val Asp Val Ser
245 250 255
       Asn Ser Pro Ile Trp Phe Gly Ile Ser Arg Het Asn Tyr Val Asn Gly
260 265 270
65
       Trp Lys Leu Ser Ser Arg Leu Asp Thr Arg Arg Gly Asp Ala Ile Gly
275 289 285
       Leu Asp Tyr Ser Gly Tyr Phe Leu Asp Thr Glu Glu Arg Ser Ser Ile
290 295 300
       Asn Leu Phe Ala Leu Leu Tyr Asn Arg Leu Arg Leu Tyr Gly Ser Trp 305 310 315
70
       His Leu Ser Asp Phe Asp Phe Ser Phe Ser Ala Asp Tyr Ala Leu Arg
325 330 335
       Gln Gly Ile Glu Arg Ile Tyr Glu Asp Tyr Lys Pro Asp Asp Asn Tyr
340 345
75
       His Ile Tyr Asp Leu Arg Ile Leu Ala Ile Arg Arg Trp Tyr Net Leu
```

```
355
                                                                           360
             Asn Glu Phe Ser Ala Gln Ala Gln Ala Ser Tyr Arg Ile Arg Thr Asp
370 375 380
             Arg Gly Cys Ala Leu Arg Val Ser Ala Gly Ser Asp Phe Tyr Gly Tyr
385 390 395
             Asp Glu Thr Tyr Arg Lys His Gl; His His Thr Het Ser Gly Met Leu
405 415
             Arg Pro Phe Ala Gly Ile Ala Tyr Asp His Ala Gly Ser Lys Leu Asp
420 430
             Phe Gly Leu Ser Leu Ser Ala Ala Tyr Arg Het Val Leu Thr His Ser
435 440
10
             Tyr Lys lie Arg Thr lie Gin Lys Glu Gin Leu Asp Tyr Gin Leu Ala
450 455 460
             Tyr Leu Pro Tyr Ala Tyr Arg Asn Arg Glu Gly Val Glu Val Arg Ser
465 470 475 486
15
             Ser Leu Tyr Val Ser Ile Pro Het Gln Asn Thr His Arg Leu Het Thr
485 490 495
             Glu Leu Arg Leu Tyr Gly Asp Leu Met Lys Arg Lys Asp Gly Ile Ala
500 505 510
             Tyr Gly Lys Thr Pro Gly Val Ile Ser His Ile Leu Ser Asp Pro Gln
515 520 525
             Ala Glu Arg Thr Ser Gly His Thr Ile Gly Ala Ile Cys Asn Ile Ser
             Tyr Leu Phe
25
              (2) INFORMATION FOR SEQ ID NO:477
                         (i) SEQUENCE CHARACTERISTICS:
30
                                    (A) LENGTH: 750 amino acids
                                     (B) TYPE: amino acid
                                    (D) TOPOLOGY: linear
                      (ii) MOLECULE TYPE: protein
                     (iii) HYPOTHETICAL: YES
                       (vi) ORIGINAL SOURCE:
                                     (A) ORGANISM: Porphyromonas gingivalis
40
                       (ix) FEATURE:
                                     (A) NAME/KEY: misc_feature
                                     (B) LOCATION 1...750
45
                       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 477
              Met Lys Lys Leu His Met Ile Ala Ala Leu Ala Val Leu Pro Phe Cys 1 \phantom{-}5\phantom{+} 10 \phantom{-}15\phantom{+}
              Leu Thr Ala Gln Ala Pro Val Ser Asn Ser Glu Ile Asp Ser Leu Ser
20 25 30
              Asn Val Gln Leu Gln Thr Val Gln Val Val Ala Thr Arg Ala Thr Ala
35 40 45
              Lys Thr Pro Val Ala Tyr Thr Asn Val Arg Lys Ala Glu Leu Ser Lys 50 60
              Ser Asn Tyr Gly Arg Asp Ile Pro Tyr Leu Leu Met Leu Thr Pro Ser 65 70 75 80
55
              Val Val Ala Thr Ser Asp Ala Gly Thr Gly Ile Gly Tyr Ser Gly Phe
85 90
              Arg Val Arg Gly Thr Asp Ala Asn Arg Ile Asn Ile Thr Thr Asn Gly 100 105 110
 60
              Val Pro Leu Asn Asp Ser Glu Ser Gln Ser Val Phe Trp Val Asn Het
115 120 125
              Pro Asp Phe Ala Ser Ser IIe Glu Asp Leu Gln Val Gln Arg Gly Val 130
Gly Thr Ser Thr Asn Gly Ala Gly Ala Phe Gly Ala Ser Val Asn Het 145
Arg Thr Asp Asn Leu Gly Leu Ala Pro Tyr Gly Arg Val Asp Leu Ser 165
Liv Cly Can Phe Gly Thr Asp Asp Asp Can Tyr Gly Arg Val Asp Leu Ser 175
Liv Cly Can Phe Gly Thr Asp Asp Asp Can Tyr Gly Arg Val Asp Can Tyr Cly Cly Can Phe Gly Thr Asp Asp Can Tyr Gly Arg Val Asp Can Tyr Cly Can Phe Gly Tyr Gly Arg Val Asp Can Tyr Cly Can Tyr Can Ty
 65
              Gly Gly Ser Phe Gly Thr Phe Arg Arg Ser Val Lys Leu Gly Ser Gly
180 180 190
 70
              Arg Ile Gly Arg His Trp Ala Val Asp Ala Arg Leu Ser Lys Ile Gly
195 200 205
               Ser Asp Gly Tyr Val Asp Arg Gly Ser Val Asp Leu Lys Ser Tyr Phe
210 215 220
               Ala Gln Val Gly Tyr Phe Gly Ser Asn Thr Ala Leu Arg Phe Ile Thr
 75
```

```
230
                                                     235
       Phe Gly Gly Lys Glu Val Thr Gly Ile Ala Trp Asn Gly Leu Ser Lys
245 250 255
       Glu Asp Glu Ala Lys Tyr Gly Arg Arg Tyr Asn Ser Ala Gly Leu Het
260 270
       Tyr Val Asp Ala Gln Gly Val Pro His Tyr Tyr His Asn Thr Asp Asn 275 280 285
       Tyr Glu Gln Arg His Tyr His Ala Ile Met Thr His Ser Phe Ser Pro
290 295 300
10
       Ser Val Ile Leu Asn Leu Thr Ala His Tyr Thr Ala Gly Tyr Gly Tyr 305 $310$
       Thr Asp Glu Tyr Arg Thr Gly Arg Lys Leu Lys Glu Tyr Ala Leu Gln 325 330 335
       Pro Tyr Val Glu Asn Ser Val Thr Val Lys Lys Thr Asp Leu Ile Arg
340 345 350
15
       Gln Lys Tyr Leu Asp Asn Asp Phe Gly Gly Leu Ile Gly Ser Leu Asn 355 360 365
       Trp His Thr Gly Ala Trp Asp Leu Gln Phe Gly Ala Ser Gly Asn Ile 370 \hspace{1cm} 375 \hspace{1cm} 380
20
       Tyr Lys Gly Asp His Phe Gly Arg Ile Thr Tyr Ile Lys Lys Tyr Ash 385 390 	 395
       Gln Pro Leu Ala Pro Asp Phe Glu Tyr Tyr Arg Asn Arg Ala Asp Lys
405 410 415
       Arg Glu Gl; Ala Ala Phe Ala Lys Ala Asn Trp Gln Ile Thr Pro Glu
420 430
25
       Leu Asn Net Tyr Ala Asp Leu Gln Tyr Arg Thr Ile Gly Tyr Thr Ile
435 440
       Asn Gly Ile Thr Asp Glu Tyr Asp Glu Val Gln Gly Ser Het Gln His
450 455 460
       Ile Asp Leu Asp Lys Thr Phe Arg Phe Leu Asn Pro Lys Ala Gly Leu 465 470 480
30
       Thr Tyr Ser Phe Asp Asp Ala His Thr Ala Tyr Ala Ser Val Ala Val
485 495
       Ala His Arg Glu Pro Asn Arg Thr Asn Tyr Thr Glu Ala Gly Ile Gly 500 505 510
35
       Gln Tyr Pro Thr Pro Glu Arg Leu Ile Asp Tyr Glu Leu Gly Tyr Arg
515 520 525
       Tyr Ala Ser Pro Leu Leu Ser Ala Gly Val Gly Leu Tyr Tyr Met Gln
530 535 540
40
       Tyr Lys Asp Gln Leu Val Leu Asp Gly Arg Leu Ser Asp Val Gly Gln
545 550 560
       Het Leu Thr Ser Asn Val Pro Asp Ser Tyr Arg Met Gly Leu Glu Leu 565 570 575
       Thr Leu Gly Trp Gln Ile Leu Pro Arg Leu Leu Arg Trp Asp Ala Ser
580 585 590
45
       Phe Thr Het Ser Arg Asn Lys lie Asp Arg Tyr Val Gln Tyr Thr Ser 595 600 605
       Val Tyr Asp Ala Asp Tyr Ash Trp Leu Glu Leu Lys Glu Glu Thr Leu
610 615
       615 626
Glu Ser Thr Asp Ile Ala Tyr Ser Pro Asn Val Ile Ala Gly Ser Het
625 630 635 640
50
       Leu Thr Leu Ser His Ala Gly Phe Glu Met Ala Trp Thr Ser Arg Phe 645 650 655
       Val Ser Lys Gln Tyr Leu Asp Asn Thr Gln Arg Ser Asp Arg Met Leu 660 665 670
55
       His Phe Val Lys Arg Val Ala Leu Gly Val Gln Leu Asn Asn Leu Phe
690 695 700
       Asn Leu Het Tyr Ala Ser Asn Ala Tyr Ile Tyr Asp Ala Gly Tyr Val
705 710 720
60
       Gln Ala Ser Gly Glu Leu Ser Ala Tyr Ala Asp Leu Arg Tyr Tyr Pro
725 730 735
       Gln Ala Gly Phe Asn Ala Leu Gly Ser Leu Thr Ile Asp Phe
65
       (2) INFORMATION FOR SEQ ID NO:478
             (i) SEQUENCE CHARACTERISTICS:
70
                   (A) LEMGTH: 494 amino acids
(B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
75
```

```
(iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                    (A) ORGANISH: Porphyromonas gingivalis
 5
            (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...494
10
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:478
       Met Lys Arg Arg Phe Leu Ser Leu Leu Leu Leu Tyr Ile Leu Ser Ser 1 5 10 15
       Ile Ser Leu Ser Ala Gln Arg Phe Pro Het Val Gln Gly Ile Glu Leu
20 25 30
15
       Asp Thr Asp Ser Leu Phe Ser Leu Pro Lys Arg Pro Trp Arg Ala Ile 35 40 45
       Gly Lys Thr Ile Gly Val Asn Leu Ala Val Trp Gly Phe Asp His Phe 50 60
       Ile Het Asn Glu Asp Phe Ala Asp Ile Ser Trp Gln Thr Ile Lys Ser 65 70 75 80
20
       Asn Phe Gln Thr Gly Phe Gly Trp Asp Asn Asp Lys Phe Val Thr Asn
85 90 95
       Leu Phe Ala His Pro Tyr His Gly Ser Leu Tyr Phe Asn Ala Ala Arg 100 $100$
25
       Ser Asn Gly Leu Ser Phe Arg His Ser Ala Pro Phe Ala Phe Phe Gly 115 120 125
       Ser Leu Het Trp Glu Leu Leu Met Glu Asn Glu Pro Pro Ser Ile Asn
130 135 140
30
       Asp Leu Cys Ala Thr Thr Ile Gly Gly Ile Ala Leu Gly Glu Met Gly
145 150 155 160
       His Arg Leu Ser Asp Leu Leu Ile Asp Asn Arg Thr Thr Gly Trp Glu
165 176 175
       Arg Met Gly Arg Glu Val Ala Ile Ala Leu Ile Asn Pro Met Arg Phe
180 185 190
35
       Leu Asn Arg Leu Thr Ala Gly Glu Val Thr Ser Val Gly Ser Arg Ser
195 200 205
       Gly Gln Ile Phe Gln Ser Val Pro Ile Asn Ile Val Val Asp Ala Gly 210 215 220
40
       Phe Arg Phe Leu Ala Asp Lys Arg His Ala Arg Thr Gly Ala Thr Ala 225 \phantom{\bigg|}230\phantom{\bigg|}230\phantom{\bigg|}235\phantom{\bigg|}
       Leu Thr Leu Asn Leu Arg Phe Asp Tyr Gly Asp Pro Phe Arg Ser Glu 245 250 255
       Thr Phe Ser Pro Tyr Asp Phe Phe Gln Phe Lys Ala Gly Leu Ser Phe 260 265 270
45
       Ser Glu Ser Gln Pro Leu Leu Ser Gln Ile Asn Leu Ile Gly Ile Leu
275 280 285
       Ser Gly Cys Gln Leu Leu Ala His Glu Arg Thr Val Leu Val Gly Gly 290 \phantom{-} 295 \phantom{-} 300
50
       Leu Phe Gln His Phe Asp Tyr Tyr Asn Ser Glu Lys Arg Ile Ser Lys 305 310 315 320
       Asn Ser Glu Glu Val Leu Val Thr Pro Tyr Arg Ile Ser Gln Val Ala
325 330 335
       Ala Leu Gly Gly Gly Leu Ile Phe Gln His His Gly Lys Phe \lambdarg Arg 340 345 350
55
       Arg Pro Leu Glu Leu Tyr Ala Glu Thr Tyr Leu Asn Val Val Pro Het 355 360 365
       Gly Ala Ser Leu Ser Asp His Tyr Asn Val Asp Asn Arg Asp Tyr Asn 370 380
60
       Leu Gly Ser Gly Leu Ser Gly Lys Leu Tyr Leu Gly Ala Thr Tyr Asn 385 390 395 400
       Asp Leu Trp Ser Trp Leu Leu Gly Val Glu Ser Tyr Arg Leu Tyr Thr 405 415
       Trp Ile Gly Tyr Glu Glu Pro His Gln Lys Asn Thr Asp Val Ser Ser 420 425 430
65
       Phe Met Val Gln Gly Asp Glu Ser Lys Ala Arg Leu Leu Val Thr Ser 435 440 445
       Ser Glu Phe Ala Phe His Pro Gly Pro Trp His Val Ala Ile Val Ala
450 455 460
70
       Arg Arg Phe Ile Arg Lys Thr Ala Tyr Gln Phe Tyr Pro Asn Val Ser
465 470 480
       Phe Asp Thr Gly Asp Ile Gln Leu Arg Val Gly Phe His Phe
75
       (2) INFORMATION FOR SEQ ID NO: 479
```

```
(i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 294 amino acids(B) TYPE: amino acid
  5
                   (D) TOPOLOGY: linear
            (ii) HOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
 10
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
 15
                  (A) MANIE/KEY: misc_feature
                  (B) LOCATION 1...294
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 479
20
       Met Lys Arg Leu Ile Val Phe Leu Ala Met Gly Gly Leu Leu Phe Thr
                                              1.0
       Leu Ala Asn Ala Gln Glu Ala Asn Thr Ala Ser Asp Thr Pro Lys Lys
       Asp Trp Thr Ile Lys Gly Val Thr Gly Leu Asn Ala Ser Gln Thr Ser
25
                35
                                    4.0
       Leu Thr Asn Trp Ala Ala Gly Gly Glu Asn Thr Val Ala Gly Asn Leu
                               55
       Tyr Leu Asn Ile Asp Ala Asn Tyr Leu Lys Asp Lys Trp Ser Trp Asp 65 70 75 80
30
       Asn Gly Leu Arg Thr Asp Phe Gly Leu Thr Tyr Thr Thr Ala Asn Lys 85 90
       Trp Asn Lys Ser Val Asp Lys Ile Glu Leu Phe Thr Lys Ala Gly Tyr 100 105 110
       Glu Ile Gly Lys His Trp Tyr Gly Ser Ala Leu Phe Thr Phe Leu Ser
115 120 125
35
       Gln Tyr Ala Lys Gly Tyr Glu Lys Pro Ser Asp His Leu Thr Gly Val
130 135
       Lys His Ile Ser Asn Phe Phe Ala Pro Ala Tyr Leu Thr Leu Gly Ile
145 150 155 160
40
       Gly Ala Asp Tyr Lys Pro Asn Glu Lys Phe Ser Leu Tyr Leu Ser Pro
165 170 175
       Thr Thr Gly Lys Leu Thr Val Val Ala Asp Asp Tyr Leu Ser Ser Leu 180 185 190
       Gly Ala Phe Gly Val Lys Val Gly Glu Lys Thr Het Phe Glu Leu Gly
195 200 205
45
       Ala Leu Val Val Gly Ser Ala Asn Ile Asn Leu Met Glu Asn Val Asn
210 215 220
       Leu Ile Thr Lys Ala Ser Phe Phe Ser Ala Tyr Thr His Asp Phe Gly
       225
                            230
                                                235
50
       Asn Ile Asp Ile Asn Trp Glu Ala Het Leu Ala Het Lys Ile Asn Lys
245 250 250
       Phe Leu Thr Ala Thr Ile Ala Thr Asn Leu Ile Tyr Asp Asp Val
260 265 270
                                        265
       Lys Ile Asn Asp Gly Pro Lys Ile Gln Phe Lys Glu Val Val Gly Val
55
               275
       Gly Val Ala Tyr Thr Phe
           290
       (2) INFORMATION FOR SEQ ID NO:480
60
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 204 amino acids
                  (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
65
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
70
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
75
                 (B) LOCATION 1...204
```

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:480
      Het Lys Lys Het Ile Leu Ala Ala Thr Met Leu Leu Ala Thr Ile Gly
 5
      Phe Ala Asn Ala Gln Ser Arg Pro Ala Leu Arg Leu Asp Ala Asn Fhe
      Val Gly Ser Asn Leu Met Gln Lys Val Ala Asn Thr Ser Val Asn Asn 35 40 45
      Lys Met Ile Val Gly Leu Arg Val Gly Ala Ala Glu Phe Ala Leu 50 ^{\circ} 60 ^{\circ}
10
      Ser Asn Asp Gly Phe Tyr Leu Ala Pro Gly Leu Ala Tyr Thr let Arg 65 75 80
      Gly Ala Lys Met Glu Ser Leu Ser Glu Thr Thr Arg Leu His Tyr
85 90 95
15
      Leu Gln Ile Pro Val Asn Ala Gly Het Arg Phe Ser Phe Ala Asp Asn 100 105 110
      Het Ala Ile Ser Leu Glu Ala Gly Pro Tyr Phe Ala Tyr Gly Val Ala 115 120 125
Gly Thr Ile Lys Thr Lys Val Ala Gly Val Thr Ala Ser Val Asp Ala 130 140
20
       Phe Gly Asp Asn Gly Tyr Asn Arg Phe Asp Leu Gly Leu Gly Leu Ser
145 150 155 160
       Ala Ala Leu Ser Tyr Asp Arg Tyr Tyr Val Gln Ile Gly Tyr Glu His
165 170 175
25
       Gly Leu Leu Asn Met Leu Lys Asp Ala Pro Asp Lys Thr Ser Leu Arg
180 185 190
       Asn His Asp Phe Phe Val Gly Leu Gly Val Arg Phe
30
       (2) INFORMATION FOR SEQ ID NO:481
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 243 amino acids
35
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
40
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
45
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...243
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:481
50
       Met Ala Ile Gly Gln Ser Arg Pro Ala Leu Arg Val Asp Ala Asn Phe
20 25 30
       Val Gly Ser Asn Gln Ser Met Lys Arg Asp Gly Tyr Val Trp Asp Thr
35 40 45
55
       Lys Het Asn Val Gly Leu Arg Val Gly Ala Ala Ala Glu Phe Het Ile
50 60
       Gly Ser Arg Gly Phe Tyr Leu Ala Pro Gly Leu Asn Tyr Thr Het Lys 65 70 75
60
       Gly Ser Lys Thr Glu Trp Asp Ile Pro Glu Het Val Pro Gly Thr Tyr
85 90 95
                        85
       Ile Thr Met Val Ser Thr Arg Leu His Tyr Leu Gln Leu Pro Ile Asn 100 \hspace{1cm} 105 \hspace{1cm} 105
65
       Ala Gly Het Arg Phe Asp Leu Het Asn Asp Het Ala Val Ser Ile Glu
115 120 125
       Ala Gly Pro Phe Leu Ala Tyr Gly Ile Tyr Gly Thr Tyr Arg Gln Lys
130 135 140
       Leu Glu Gly Trp Lys Pro Asn Asn Tyr Ser Thr Glu Phe Phe Gly Pro
145 155 160
 70
        Thr Leu Gly Gly Pro Thr Asn Ile Arg Trp Asp Ile Gly Ala Asn Ile
165 170 175
        Ile Ala Ala Phe His Tyr Lys Arg Tyr Tyr Ile Gln Ile Gl; Tyr Glu
180 185 190
        His Gly Phe Val Asp Ile Val Ser Gly Gly Gly Ser Asp Ile Pro Arg
 75
```

```
195
                                       200
                                                             205
      Leu Asn Asp Asn Arg Gln Ser Ser Ser Thr Thr Ala Leu Arg Glu Lys
210 215 220
                                                     220
                              215
           210
      Gly Asn Asn Glu Tyr Ala Tyr Asn Arg Asp Phe Phe Val Gly Ile Gly
      Tyr Arg Phe
       (2) INFORMATION FOR SEQ ID NO:482
10
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 207 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
15
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
20
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
25
                  (B) LOCATION 1...2\overline{07}
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:482
      Met Lys Arg Net Leu Leu Leu Val Val Leu Leu Tyr Gly Ile Ala
30
      Gly Arg Leu Ala Ala Gln Asp Val Ile Arg Pro Trp Ser Leu Gln Val 20 \\ 0 \\ 25 \\ 30
      Gly Ala Gly Tyr Ser Asp Thr Glu Asn Ile Pro Gly Gly Phe Thr Tyr 35 40 45
35
      Gly Phe Tyr Leu Gly Lys Arg Het Gly Ser Phe Leu Glu Val Gly Leu 50 60
      Ser Het Tyr Asn Ser Thr Arg Gln Thr Ala Asn Asn Ala Asp Ser Phe 65 70 75 80
      Ala Ser Asn Glu Gly Asp Gly Ser Phe Gln Val Asn Met Ser Ser Pro
85 90 95
40
      Asn Glu Lys Trp Ser Phe Phe Asp Ala Gly Ser Ala Asn Cys Tyr Het 100 \, 105 \, 110
       Ile Val Val Gl; Val Asn Pro Leu His Leu Phe Trp Gln Asn Ser Arg
115 120 125
      His Asn Leu Phe Leu Ala Val Gln Ala Gly Leu Ser Asn Lys His Asn 130 135 140

Ile His Fhe Ile Tyr Gly Asp Lys Gly Ala Lys Val Ser Ile Tyr Thr 145 150 150 165
45
      Asn Ser Asn Thr Tyr Ile Gly Tyr Gly Ala Arg Val Ala Tyr Glu Tyr
165 170 175
50
       Gln Ile His Lys Asn Val Gly Ala Gly Ala Ala Val Het Tyr Asp His
       Gly Asn Lys Net Leu Thr Ala Met Ala Thr Leu Ser Thr His Phe
55
       (2) INFORMATION FOR SEQ ID NO:483
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 951 amino acids
60
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
65
           (ili) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
70
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...951
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:483
75
```

| | 1 | | | | 5 | | | Ser | | 10 | | | | | 12 | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | lle | | | 20 | | | | Ala | 25 | | | | | 30 | | |
| 5 | | | 3 5. | Ala | | | | His 40 | | | | | 45 | | | |
| | | 50 | | | | | 55 | Val | | | | ゆり | | | | |
| 10 | 65 | Alā | | | | 70 | | Pro | | | 75 | | | | | 80 |
| 10 | Gln | | | | 85 | | | Glu | | 90 | | | | | 95 | |
| | | | | 100 | | | | Val | 105 | | | | | 110 | | |
| 15 | | | 115 | | | | | Ser 120 | | | | | 125 | | | |
| | | 130 | Het | | | | 135 | Asp | | | | 140 | | | | |
| 20 | 145 | Leu | | | | 150 | | Pro | | | 155 | | | | | T-00 |
| | | | | | 165 | | | Lys | | 170 | | | | | 1/5 | |
| | | | | 180 | | | | Vāl | 185 | | | | | 190 | | |
| 25 | | | 195 | | | | | Lys 200 | | | | | 205 | | | |
| | | 210 | | | | | 215 | Сле | | | | 220 | | | | |
| 30 | Tyr 225 | Ile | Asn | Leu | Ser | Lys 230 | Ala | Lys | Val | His | Pro 235 | Glu | Lys | Asp | Ile | Val 240 |
| 00 | Thr | G1?. | Pro | Val | Asn 245 | Leu | Val | Ile | Alā | Asp 250 | | Pro | Leu | Pro | 11e 255 | Gl ; |
| | Leu | Pro | Phe | Gly 260 | | Phe | Pro | Phe | Ser 265 | | Lys | T)r | Ser | Ser 270 | | Ile |
| 35 | Leu | Het | Pro 275 | Thr | Туг | Gly | Glu | Asp 280 | Asn | Arg | Tyr | Gly | Phe 285 | Tyr | Leu | Arg |
| | Asn | Gly 290 | Gly | Tyr | Tyr | Phe | Ala 295 | Phe | Ser | Asp | Туг | 11e 300 | Asp | Leu | Ala | Leu |
| 40 | Arg 305 | G17. | Glu | Ile | Phe | Ser 310 | Lys | Gly | Ser | Trp | Glչ 315 | Ile | Ser | Ala | Gln | Ser 320 |
| 10 | r?.e | | | | 325 | Tyr | | Туr | | 330 | | | | | 335 | |
| | | | | 340 | Ser | | | Lγε | 345 | | | | | 350 | | |
| 45 | | | 355 | | | | | Trp 360 | | | | | 365 | | | |
| | | 370 | | | | | 375 | | | | | 380 | | | | |
| 50 | 385 | | | | | 390 | | Thr | | | 395 | | | | | 400 |
| | | | | | 405 | | | Val | | 410 | | | | | 415 | |
| | | | | 420 | | | | | 425 | | | | | 430 | | Arg |
| 55 | | | 435 | | | | | 440 | | | | | 445 | | | Ser |
| | | 450 |) | | | | 455 | | | | | 460 | | | | Trp |
| 60 | 465 | , | | | | 470 | | | | | 475 | 1 | | | | Ile 480 |
| | | | | | 485 | | | | | 490 | | | | | 495 | |
| | | | | 500 |) | | | | 505 | , | | | | 510 | | Leu |
| 65 | | | 51.5 | 5 | | | | 520 |) | | | | 525 | • | | Trp |
| | | 530 |) | | | | 535 | , | | | | 540 |) | | | Phe |
| 70 | 545 | 5 | | | | 550 |) | | | | 555 | , | | | | Ser 560 |
| | Lei | ı Sei | | | 565 | , | | | | 570 |) | | | | 575 | |
| | | | | 580 |) | | | | 585 | 5 | | | | 590 |) |) Phe |
| 7 5 | Thi | r Pro | Thi | | | Phe | e Sei | ту1 | | | Asp | Phe | Th: | : Lys | Arç | , Arg |

```
595
                                          600
                                                                 605
        Tyr Gly Phe Trp Glu Leu Leu Glu His Thr Asp Gln Asn Gly Lys Leu 610 620
        His Thr Leu Leu Tyr Ser Pro Tyr Phe Glu Gln Ile Phe Gly Ala Pro 625 630 635 640
  5
        Ser Het Gly Asn Ala Gly Ser Val Asn Phe Ser Phe Asp Asn Asn Leu 645 650 655
        Glu Ala Lys Ile Lys Ser Lys Ser Asp Ser Thr Gly Ile Lys Lys Ile
660 665 670
 10
        Ser Leu Ile Asp Gln Phe Thr Trp Ser Thr Ser Tyr Asn Met Phe Ala
675 680 685
        Asp Ser Ile Arg Trp Ser Asm Ile Ser Ala Ser Leu Ala Leu Arg Leu
690 695 700
        Ser Lys Ser Phe Thr Leu Arg Leu Ser Gly Leu Fre Asp Pro Tyr Leu 705 715 720

Thr Lys Tyr Tyr Glu Gly Glu Asp Gly Lys Ile Ile Pro Tyr Lys Ser 725 730 735
 15
        Asn Asp Leu Arg Ile Phe Asn Gly Lys Gl; Leu Ala Arg Leu Ile Ser
740 745 750
20
        Thr Gly Thr Ser Phe Ser Tyr Thr Leu Asn Lys Glu Ser Leu Ser Gly 755 760 765
        Leu Île Ala Leu Phe Ser Gly Lys Lys Glu Arg Arg Asp Glu Lys Lys
770 780
        Asn Thr Gly Ala Thr Pro His Glu Gly Asp Asp Ala Ala Asp Ile Leu
785 790 795 800
25
        Glu Gly Gly Arg Pro Gln Asn Glu Ser Gly Gly Ser Leu Leu Glu Arg
805 810 815
        Asn Arg Gln Gly Gly Ala Val Asp Gln Asp Gly Tyr Phe Ala Tyr Ser
820 825 830
30
        Ile Pro Trp Ser Leu Ser Phe Asp Tyr Ser Trp Asn Ile Ala Thr Asp
835 840 845
        Tyr Asn Arg Tyr Asn Val Asn Lys Met Glu His Tyr Tyr Arg Val Thr
850 855 860
        Gln Asn Leu Ser Phe Arg Gly Asn Ile Gln Pro Thr Pro Asn Trp Ser
865 870 875 880
35
        Phe Gly Phe Asn Ala Asn Tyr Asn Phe Asp Leu Lys Lys Ile Thr Ser
885 890 895
        Leu Thr Cys Asn Val Thr Arg Asp Met His Cys Trp Ala Ile Ser Ala
900 905 910
        Ser Phe Ile Pro Ile Gly Ala Tyr Lys Ser Tyr Asn Phe Val Ile Ser
915 920 925
40
        Val Lys Ser Ser Leu Leu Gln Asp Leu Lys Tyr Gln Gln Ser Asn Arg
930 935 940
        Pro Ile Thr Asn Thr Trp Tyr
45
        945
        (2) INFORMATION FOR SEQ ID NO: 484
              (i) SEQUENCE CHARACTERISTICS:
50
                   (A) LENGTH: 1226 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
55
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...1\overline{2}26
65
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:484
       Het Het Lys Arg Tyr Thr He He Leu Ala Val Phe Leu Leu Phe Cys
       Thr Val Phe Thr Phe Gln Ile Lys Ala Arg Fro Tyr Glu Arg Phe Ala
20 25 30
70
       Asp Val Glu Lys Pro Trp Ile Gln Lys His Ser Het Asp Ser Lys Leu
35 40
       Val Pro Ala Asn Lys Gly Asn Leu Ile Gln Ala Glu Ile Val Tyr Gln
50 60
75
       Ser Val Ser Glu His Ser Asp Leu Val Ile Ser Pro Val Asn Glu Ile
```

| | | | | | | 74 | | | | | 76 | | | | | 80 |
|----|------|------------|------------|------------|-----|-----|------------|------------|-------------|------|------------|------------|------------|------------|-------|------------|
| | | | | | 85 | | | | | 90 | 75 Lys | | | | 95 | Glu |
| 5 | | | | 100 | | | | | 105 | | Val | | | 110 | | |
| | | | 115 | | | | | 120 | | | Glu | | 125 | | | |
| | | 1.30 | | | | | 135 | | | | Val | 140 | | | | |
| 10 | 145 | Val | | | | 150 | | | | | Het 155 | | | | | 160 |
| | Ala | | | | 165 | Ala | | | | 170 | Gln | | | | 175 | |
| 15 | Het | Leu | Pro | Lys 180 | Gly | Thr | Alā | Thr | Gl u 185 | Gl y | Vāl | Arg | Ile | Ala 190 | Leu | Väl |
| | Ser | T;r | Asp 195 | His | Glu | Pro | His | Arg 200 | Leu | Ser | Asp | Phe | Thr 205 | Γλε | Asp | Thr |
| | Al.ā | Phe 210 | Leu | C2.2 | Gln | L∵s | 11e 215 | Arg | Ala | Leu | Thr | Pro 220 | Il∈ | Тгр | G1 λ. | Thr |
| 20 | 225 | Thr | | | | 230 | Γ.λ.ε | | | | Asn 235 | | | | | 240 |
| | Thr | | | | 245 | His | | | | 250 | Ser | | | | 255 | |
| 25 | | | | 260 | Val | | | | 265 | | Ala | | | 270 | | |
| | Thr | Gly | Asn 275 | Ala | Asn | Asp | Pro | Ile 280 | Asp | Leu | Val | Ile | Gln 285 | Gly | Ala | Ile |
| | Asn | Phe 290 | Pro | Thr | Asn | Tyr | Val 295 | Ser | Asn | Asn | Pro | Ser 300 | Thr | Pro | Leu | Thr |
| 30 | 305 | | | | | 310 | | | | | Gly 315 | | | | | 320 |
| | G] u | | | | 325 | | | | | 330 | Ala | | | | 335 | |
| 35 | | | | 340 | | | | | 345 | | Arg | | | 350 | | |
| | | | 355 | | | | | 360 | | | Asn | | 365 | | | |
| | | 370 | | | | | 375 | | | | Gl y | 380 | | | | |
| 40 | 385 | | | | | 390 | | | | | Thr 395 | | | | | 400 |
| | | | | | 405 | | | | | 410 | Ala | | | | 415 | |
| 45 | | | | 420 | | | | | 425 | | Arg | | | 430 | | |
| | | | 435 | | | | | 440 | | | Leu | | 445 | | | |
| | | 450 | | | | | 455 | | | | Thr | 460 | | | | |
| 50 | 465 | | | | | 470 | | | | | Thr 475 | | | | | 480 |
| | | | | | 485 | | | | | 490 | | | | | 495 | |
| 55 | | | | 500 | | | | | 505 | | Ala | | | 510 | | |
| | | | 515 | | | | | 520 | | | | | 525 | | | Ser |
| | | 530 | | | | | 535 | | | | | 540 | | | | r/.a |
| 60 | 545 | | | | | 550 | | | | | 555 | | | | | Asp 560 |
| | | | | | 565 | | | | | 570 | | | | | 575 | |
| 65 | | | | 580 | 1 | | | | 585 | | | | | 590 | | Gly |
| | | | 595 | 1 | | | | 600 | 1 | | | | 605 | 1 | | Leu |
| 70 | | 610 |) | | | | 615 | , | | | | 620 | ı | | | Asp |
| 70 | 625 | , | | | | 630 | | | | | 635 | , | | | | Pro 640 |
| | | | | | 645 | II. | | | | 650 |) | | | | 655 | |
| 75 | Ala | T;r | Lys | 660 | | Aan | туг | Trp | Het 665 | | , Gl | Thr | Thr | 670 | GIN | Gln |

```
Thr Gly Glu Asp Val Glu Phe Ala Thr Thr Glu Asn Phe Gly Ser Pro 690 700
       Ala Val Ala Asp Leu His Val Pro Thr Thr Asn Pro Lys Ile Ile Gly 705 710 715 720
       Asn Leu Ile Asn Asn Ser Asp Lys Asp Leu Val Val Thr Thr Ser Ser
725 730 735
       Gln Leu Thr Ile Asn Gly Val Val Glu Asp Asn Asn Pro Asn Val Gly 740 745 750
10
       Thr Ile Val Val Lys Ser Ser Lys Asp Asn Pro Thr Gly Thr Leu Leu
755 760 765
       Phe Ala Asn Pro Gly Tyr Asn Gln Asn Val Gly Gly Thr Val Glu Phe
770 775 790
       Tyr Asn Gln Gly Tyr Asp Cys Ala Asp Cys Gly Het Tyr Arg Arg Ser
785 790 795 800
15
       Trp Gln Tyr Fhe Gly Ile Pro Val Asn Glu Ser Gly Phe Pro Ile Asn 805 816
       Asp Val Gly Gl; Asn Glu Thr Val Asn Gln Trp Val Glu Pro Phe Asn 820 825 830
20
       Gl; Asp L;s Trp Arg Pro Ala Pro T;r Ala Pro Asp Thr Glu Leu Gln
835 840 845
       Lys Phe Lys Gly Tyr Gln Ile Thr Asn Asp Val Gln Ala Gln Pro Thr
850 855 860
       Gly Val Tyr Ser Phe Lys Gly Het Ile Cys Val Cys Asp Ala Phe Leu
865 870 870 880
25
       Asn Leu Thr Arg Thr Ser Gly Val Asn Tyr Ser Gly Ala Asn Leu Ile
885 890 895
       Gly Asn Ser Tyr Thr Gly Ala Ile Asp Ile Lys Gln Gly Ile Val Phe 900 905
30
       Pro Pro Glu Val Glu Gln Thr Val Tyr Leu Phe Asn Thr Gly Thr Arg
915 920 925
       Asp Gln Trp Arg Lys Leu Asn Gly Ser Thr Val Ser Gly Tyr Arg Ala
930 935 940
       Gly Gln Tyr Leu Ser Val Pro Lys Asn Thr Ala Gly Gln Asp Asn Leu
945 950 950 960
35
       Pro Asp Arg Ile Pro Ser Met His Ser Phe Leu Val Lys Met Gln Asn 965 970 975
       Gly Ala Ser Cys Thr Leu Unk Ile Leu Tyr Asp Lys Leu Leu Lys Asn
980 985 990
40
       Thr Thr Val Asn Asn Gly Asn Gly Thr Gln Ile Thr Trp Arg Ser Gly 995 1000 1005
       Asn Ser Gly Ser Ala Asn Met Pro Ser Leu Val Met Asp Val Leu Gly
                     1015
         1010
                                                        1020
45
       Asn Glu Ser Ala Asp Arg Leu Trp Ile Phe Thr Asp Gly Gly Leu Ser 1025 1030 1035
       Phe Gly Phe Asp Asp Gly Trp Asp Gly Arg Lys Leu Thr Glu Lys Gly
1045 1050 1055
       Leu Ser Gln Leu Tyr Ala Het Ser Asp Ile Gly Asn Asp Lys Phe Gln
1060 1065 1070
50
       Val Ala Gly Val Pro Glu Leu Asn Asn Leu Leu Ile Gly Phe Asp Ala
1075 1080 1085
       Asp Lys Asp Gly Gln Tyr Thr Leu Glu Phe Ala Leu Ser Asp His Phe
1090 1095 1100
       Ala Lys Gly Ala Val Tyr Leu His Asp Leu Gln Ser Gly Ala Lys His
1105 1110 1115
55
       Arg Ile Thr Asn Ser Thr Ser Tyr Ser Phe Asp Ala Lys Arg Gly Asp 1125 1130 1135
       Ser Gly Ala Arg Phe Arg Leu Ser Tyr Gly Cys Asp Glu Asn Val Asp
1140 1145
60
       Asp Ser His Val Val Ser Thr Asn Gly Arg Glu Ile Ile Ile Leu Asn
1155 1160 1165
       Gln Asp Ala Leu Asp Cys Thr Val Thr Leu Phe Thr Ile Glu Gly Lys
1170 1175 1180
       Leu Leu Arg Arg Leu Lys Val Leu Ala Gly His Arg Glu Val Het Lys
1185 1190 1195 1200
65
       Val Gln Thr Gly Gly Ala Tyr Ile Val His Leu Gln Asn Ala Phe Thr
1205 1210 1215
                         1205
                                               1210
       Asn Asp Val His Lys Val Leu Val Glu Tyr
70
       (2) INFORMATION FOR SEQ ID NO:485
```

(i) SEQUENCE CHARACTERISTICS:

75 (A) LENGTH: 1225 amino acids

```
(B) TYPE: amino acid
                   (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
 5
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
10
            (ix) FEATURE:
                   (A) NAHE/KET: misc_feature
                   (B) LOCATION 1...1225
15
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 485
       Het Lys Arg Tyr Thr Ile Ile Leu Ala Val Phe Leu Leu Phe Cys Thr
       Val Phe Thr Phe Gin Ile Lys Ala Arg Pro Tyr Glu Arg Phe Ala Asp
20 25 30
20
       Val Glu Lys Pro Trp Ile Gln Lys His Ser Het Asp Ser Lys Leu Val
35 40
       Pro Ala Asn Lys Gly Asn Leu Ile Gln Ala Glu Ile Val Tyr Gln Ser 50 60
       Val Ser Glu His Ser Asp Leu Val Ile Ser Pro Val Asn Glu Ile Arg
65 70 80
25
       Pro Ala Ash Arg Phe Pro Ser His Arg Lys Ser Phe Phe Ala Glu Ash
85 90 95
       Leu Arg Ala Ser Pro Pro Val Val Pro Val Ala Val Asp Lys Tyr Ala 100 105 110
30
       Val Pro Val Ala Asn Pro Met Asp Pro Glu Asn Pro Asn Ala Trp Asp
115 120 125
       Val Thr Leu Lys Ile Thr Thr Lys Ala Val Thr Val Pro Val Asp Val
139 135 140
       Val Met Val Ile Asp Gln Ser Ser Ser Met Gly Gly Gln Asn Ile Ala
145 150 155 160
35
       Arg Leu Lys Ser Ala Ile Ala Ser Gly Gln Arg Phe Val Lys Lys Met
165 170 175
       Leu Pro Lys Gly Thr Ala Thr Glu Gly Val Arg Ile Ala Leu Val Ser
40
       Tyr Asp His Glu Pro His Arg Leu Ser Asp Phe Thr Lys Asp Thr Ala
195 200 205
       Phe Leu Cys Gln Lys Ile Arg Ala Leu Thr Pro Ile Trp Gly Thr His
210 215 220
       Thr Gln Gly Gly Leu Lys Met Ala Arg Asn Ile Met Ala Thr Ser Thr 225 230 235 240
45
       Ala Val Asp Lys His Ile Ile Leu Met Ser Asp Gly Leu Ala Thr Glu
245 250 255
       Gln Tyr Pro Val Lys Asn Val Thr Thr Ala Asp Phe Ile Gly Lys Thr
260 265 270
50
       Gly Asn Ala Asn Asp Pro Ile Asp Leu Val Ile Gln Gly Ala Ile Asn 275 280 285
        Phe Pro Thr Asn Tyr Val Ser Asn Asn Pro Ser Thr Pro Leu Thr Pro
290 295 300
       Asn Tyr Pro Thr His Ser Ser Lys Val Gly Arg Arg Asn Leu Pro Glu
305 310 315
55
       Ser Lys Phe Asp Tyr Ser Asn Leu Ser Ala Arg Ile Thr Phe Asp Gly 325 330 335
       Val Ala Gly Ala Leu Val Tyr Glu Pro Arg Phe Pro His Pro Tyr Tyr 340 345 350

Tyr Tyr Phe Pro Cys Asn Ala Ala Ile Asn Glu Ala Gln Phe Ala Lys 355 360 360
60
        Asn Ser Gly Tyr Thr Ile His Thr Ile Gly Tyr Asp Leu Gly Asp Phe
370 375 380
        Ala Leu Ala Asn Asn Ser Leu Lys Leu Thr Ala Thr Asp Glu Asn His
385 390 395 400
65
        Phe Phe Thr Ala Thr Pro Ala Asn Leu Ala Ala Ala Phe Asp Asn Ile
405 410 415
        Ala Gln Thr Ile Asn Ile Gly Ile Gln Arg Gly Glu Val Thr Asp Phe 420 425 430
 70
        Val Ala Pro Gly Phe Ile Val Lys Asn Leu Thr Gln Ser Gly Asp Val
435 440 445
        Thr His Leu Leu Asn Val Ser Asn Gly Thr Val His Tyr Asp Val Ser
                                 455
                                                       460
 75
        Thr Lys Lys Leu Thr Trp Thr Thr Gly Thr Ile Leu Ser Ser Ser Glu
```

| | 465 | | | | | 470 | | | | | 475 | | | | | 480 |
|------------|-------|------------|------------|------------|------------|-----|------------|------------|-------------|------------|-----|------------|------------|------------|------------|------|
| | Ala | Thr | Ile | Thr | Tyr 485 | Arg | Ile | Tyr | Ala | Asp 490 | Leu | Asp | Tyr | Ile | Gln 495 | Asn |
| 5 | Asn | Asp | Il∈ | Pro 500 | | Asn | Thr | Thr | Ser 505 | | Ile | Gly | Pro | Asp 510 | | GlŢ |
| - | Gly | Phe | Asp 515 | Thr | Asn | Thr | Glu | Ala 520 | Lys | Leu | Thr | туг | Thr 525 | Asn | Ser | Asn |
| | Gl y | Glu 530 | Pro | Asn | Gln | Gln | Leu 535 | Ile | Phe | Pro | Arg | Pro 540 | Thr | Val | Lys | Leu |
| 10 | 545 | _ | Gl y | | | 550 | | | _ | | 555 | | | | | 560 |
| | | | Ile | | 565 | | | | | 570 | | | | | 575 | |
| 15 | | | Leu | 580 | | | - | | 585 | | | | - | 590 | _ | |
| | | | 9ro 595 | | | | | 600 | - | | | | 605 | | | |
| 00 | - | 610 | Ser | | | | 615 | | | | | 620 | | | | |
| 20 | 625 | | T'.r | | | 630 | | | | | 635 | | | | | 640 |
| | | | Gly | | 645 | | _ | | | 650 | | | | | 655 | |
| 25 | T7.1. | Lys | Leu | 660 Fen | Asn | Tyr | Trp | Met | G1 y 665 | G17 | Thr | Thr | Asp | 670 | Gin | Ser |
| | Glu | Trp | Asp 675 | Val | Thr | Ser | Asn | Trp 680 | Thr | G1y | Λla | Gln | Val 685 | Pro | Leu | Thr |
| 00 | | 690 | Asp | | | | 695 | | | | | 700 | | | | |
| 30 | 705 | | Asp | | | 710 | | | | | 715 | | | | | 720 |
| | | | Asn | | 725 | | _ | - | | 730 | | | | | 735 | |
| 35 | | | Ile | 740 | | | | | 745 | | | | | 750 | | |
| | | | Val 755 | | | | | 760 | | | | | 765 | | | |
| 40 | | 770 | Pro | | | | 775 | | | | | 780 | | | | |
| 40 | 785 | | Gly | | | 790 | | | | | 795 | | | | | 800 |
| | | | Phe | _ | 805 | | | | | 810 | | | | | 815 | |
| 4 5 | | | Gly | 820 | | | | | 825 | | | | | 830 | | |
| | | | Trp 835 | | | | | 840 | | | | | 845 | | | |
| EΩ | | 850 | Gly | _ | | | 855 | | | | | 860 | | | | - |
| 50 | 865 | | Ser | | | 870 | | | | | 875 | | | | | 880 |
| | | | Arg | | 885 | _ | | | _ | 890 | _ | | | | 895 | _ |
| 55 | | | Tyr | 900 | _ | | | = | 905 | - | | _ | | 910 | | |
| | | | Val 9)5 | | | | | 920 | | | | | 925 | | | |
| co | | 930 | Arg | - | | | 935 | | | | | 940 | | _ | | |
| 60 | 945 | | Leu | | | 950 | | | | | 955 | | | | | 960 |
| | | | Ile | | 965 | | | | | 970 | | | | | 975 | |
| 65 | | | C7.2 | 980 | | | | | 985 | _ | - | | | 990 | | |
| | | | Asn 995 | | | | | 100 | 0 | | | | 10 | 05 | | |
| 70 | | 101 | | | | | 10 | 15 | | | | 1 | 020 | | _ | |
| 70 | 102 | 5 | Ala | _ | _ | 103 | 30 | | | | 1 | 035 | _ | | | 1040 |
| | | | Asp | | 104 | 5 | | | | 10 | 50 | | | | 1 | 055 |
| 75 | ser | GID | Leu | Туг 106 | | Het | Ser | Asp | lle 10 | | Asn | Asp | rlæ | | Gln 070 | Val |

```
Ala Gly Val Pro Glu Leu Asn Asn Leu Leu Ile Gly Phe Asp Ala Asp
1075 1080 1085
       Lys Asp Gly Gln Tyr Thr Leu Glu Phe Ala Leu Ser Asp His Phe Ala
                                                             1100
                                1095
           1090
       Lys Gly Ala Val Tyr Leu His Asp Leu Gln Ser Gly Ala Lys His Arg
1105 1110 1115
       Ile Thr Asn Ser Thr Ser Tyr Ser Phe Asp Ala Lys Arg Gly Asp Ser
1125 1130 1135
       Gly Ala Arg Phe Arg Leu Ser Tyr Gly Cys Asp Glu Asn Val Asp Asp
1140 1145
10
       Ser His Val Val Ser Thr Asn Gly Arg Glu Ile Ile Leu Asn Gln
1155 1160 1165
       Asp Ala Leu Asp Cys Thr Val Thr Leu Phe Thr Ile Glu Gly Lys Leu
1170 1175 1180
       Leu Arg Arg Leu Lys Val Leu Ala Gly His Arg Glu Val Met Lys Val
1185 1190 1195
15
       Gln Thr Gl; Gly Ala Tyr Ile Val His Leu Gln Asn Ala Phe Thr Asn
1205 1210 1215
       Asp Val His Lys Val Leu Val Glu Tyr
20
                     1220
                                               1225
        (2) INFORMATION FOR SEQ ID NO:486
              (i) SEQUENCE CHARACTERISTICS:
25
                    (A) LENGTH: 425 amino acids
                    (P) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
30
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
35
             (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...425
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:486
40
        Het Glu Val Lys Lys Asn Thr Val Val Leu Arg Leu Leu Ile Trp Phe
        Val Ala Ile Leu Leu Phe His Ser Ser Arg Leu Trp Gly Gln Glu Gly 20 25 30
45
        Glu Gly Ser Ala Arg Tyr Arg Phe Lys Gly Phe Val Asp Thr Tyr His
35 40 45
        Ala Val Arg Ser Ser Ser Pro Phe Asp Phe Met Ser Ser Arg Thr Arg
50 55 60
        Val Arg Gly Glu Leu Glu Arg Ser Phe Gly Asn Ser Lys Val Ala Val
65 70 75 80
50
        Ser Val Asn Ala Thr Tyr Asn Ala Leu Leu Lys Asp Glu Thr Gly Leu
85 90 95
        Arg Leu Arg Glu Ala Phe Phe Glu His Gln Glu Glu His Trp Gly Leu
100 105 110
55
        Arg Leu Gly Arg Gln Ile Val Ile Trp Gly Ala Ala Asp Gly Val Arg
115 120 125
        Ile Thr Amp Leu Ile Ser Pro Met Amp Met Thr Glu Phe Leu Ala Gln
130 140
        Asp Tyr Asp Asp Ile Arg Met Pro Val Asn Ala Leu Arg Phe Ser Val 145 150 150 166

Phe Asn Glu Ser Met Lys Val Glu Val Val Val Leu Pro Val Phe Glu 175 165 175 175
60
        Gly Tyr Arg Leu Pro Val Asp Pro Arg Asn Pro Trp Asn Ile Phe Ser
180 185 190
 65
        Leu Ser Pro Ile Ala Gln Gly Met Asn Ile Val Trp Lys Glu Glu Ala
195 200 205

Gly Lys Pro Ala Phe Lys Val Ala Asn Ile Glu Tyr Gly Ala Arg Trp
210 215 220
         Ser Thr Thr Leu Ser Gly Ile Asp Phe Ala Leu Ala Ala Leu His Thr
225 230 235
 70
        Trp Asn Lys Met Pro Val Ile Glu Val Gln Gly Ile Val Pro Thr Glu 245 250 255
         Ile Ile Val Ser Pro Arg Tyr Tyr Arg Het Gly Phe Val Gly Gly Asp 260 270
 75
```

```
Leu Ser Val Pro Val Gly Gln Phe Val Phe Arg Gly Glu Ala Ala Phe
                                      280
                                                           285
       Asn Ile Asp Lys His Phe Thr Tyr Lys Ser His Ala Glu Gln Glu Gly
290 295 300
       Phe Gln Thr Ile Asn Trp Leu Ala Gly Ala Asp Trp Tyr Ala Pro Gly 305 310 315 320
       Glu Trp Het Ile Ser Gly Gln Phe Ser Het Glu Ser Ile Phe Arg Tyr 325 330 335
                                                                 335
       Arg Asp Phe Ile Ser Gln Arg Gln His Ser Thr Leu Ile Thr Leu Asn 340 345 350
 10
       Val Ser Lys Lys Phe Phe Gly Ser Thr Leu Gln Leu Ser Asp Phe Thr
355 360 365
       Tyr Tyr Asp Leu Thr Gly Lys Gly Trp Phe Ser Arg Phe Ala Ala Asp 370 375
       Tyr Ala Leu Asn Asp Gln Ile His Leu Met Ala Gly Tyr Asp Trp Phe
385 390 395 400
15
       Ser Ser Lys Gly Ser Gly Ile Phe Asp Arg Tyr Lys Asp Asn Ser Glu
405 410 415
       Leu Trp Phe Lys Ala Arg Tyr Ser Phe
20
                   420
       (2) INFORMATION FOR SEQ ID NO:487
             (i) SEQUENCE CHARACTERISTICS:
25
                  (A) LENGTH: 404 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
30
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...404
40
           (xi) SEQUENCE DESCRIPTION: SEO ID NO:487
      Het Ser Ser Cys Glu Val Ala Tyr Phe Ser Leu Lys Pro Ile Asp Leu 1 	ext{ } 5 	ext{ } 10 	ext{ } 15
       Gln Asn Ile Arg Glu Arg Asn His Ser Ser Asp Ile Ala Leu Ser Asn
20 25 30
45
       Leu Leu Asp Asn Ser Asn Gln Leu Leu Ala Thr Ile Leu Ile Gly Asn 35 40 45
       Asn Val Ile Asn Val Ala Ile Val Ile Leu Ser Asn Tyr Ala Ile Glu
50 55 60
50
       Gln Thr Phe Val Phe Ser Ser Pro Ile Ile Gly Phe Leu Ile Gln Thr
65 70 75 80
      Ile Leu Leu Thr Thr Val Leu Leu Leu Phe Gly Glu Ile Leu Pro Lys 85 90 95
      Val Tyr Ala Arg Lys Asn Pro Leu Gln Tyr Ser Arg Phe Ser Ala Ala
100 105 110
55
      Ala Het Ser Val Ile Tyr Lys Ile Leu Ser Pro Phe Ser Lys Leu Leu 115 120 125
      Val Lys Ser Thr Gly Ile Val Thr Arg Gly Ile Ser Lys Lys Tyr
130 135 140
60
      Asp Het Ser Val Asp Glu Leu Ser Lys Ala Val Ala Leu Thr Thr Thr 145 150 160
      Glu Gly Glu Pro Glu Glu Lys Glu Met Ile Asn Glu Ile Ile Lys Phe
165 170 175
      Tyr Asn Lys Thr Ala Cys Glu Ile Met Val Pro Arg Ile Asp Ile Val
65
      70
      Ile Lys Gly Val 11e Tyr Ile Lys Asp Leu Ile Pro His Net Asp Lys225230235240
      Gly Asp Glu Phe Asp Trp His Pro Leu Ile Arg Lys Ala Tyr Phe Val
245 250 255
      Pro Glu Asn Lys Arg Ile Asp Asp Leu Leu Glu Glu Phe Arg Ala Asn
75
```

```
Lys Val His Val Ser Ile Val Val Asp Glu Phe Gly Gly Thr Cys Gly 275 280 285
      Leu Ile Thr Met Glu Asp Ile Leu Glu Glu Ile Val Gly Glu Ile Thr
                                295
         290
      Asp Glu Tyr Asp Glu Glu Glu Leu Pro Phe Lys Val Leu Gly Asp Gly 305 310 315
      Ser Tyr Leu Phe Glu Gly Lys Thr Ser Leu Ser Asp Val Arg His Tyr
325 335
      Leu Asp Leu Pro Glu Asn Ala Phe Gly Glu Leu Gly Asp Glu Val Asp 340 350
10
      Thr Leu Ser Gly Leu Phe Leu Glu Ile Lys Gln Glu Leu Pro His Val
355 360 365
      Gly Asp Thr Ala Val Tyr Glu Pro Phe Arg Phe Gln Val Thr Gln Het 370 375 380
      Asp Lys Arg Arg Ile Ile Glu Ile Lys Ile Phe Pro Phe Glu Arg Thr
15
      Trp Glu Val Glu
20
       (2) INFORMATION FOR SEQ ID NO:488
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 260 amino acids
                  (B) TYPE: amino acid
25
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
30
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...260
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:488
       Met Lys Leu Leu Tyr Leu Leu Leu Val Leu Ser Thr Leu Ser Pro
40
                                              10
       Het Tyr Ser Gln Met Leu Phe Ser Glu Asn Leu Thr Het Asn Ile Asp
20 25 30
       Ser Thr Lys Thr Ile Gln Gly Thr Ile Leu Pro Val Leu Asp Phe Lys 35 40 45
45
       Thr Glu Lys Glu Asn Val Phe Thr Phe Lys Asn Thr Ala Asn Leu Asn 50 55 60
       Leu Leu Ile Lys His Gly Gin Val Ile Asn Leu Ile Asn Lys Leu Glu
65 70 80
       Phe Ser Thr Tyr Gly Asn Lys Val Thr Val Ser Gly Gly Tyr Val His
85 90 95
50
       Thr Glu Tyr Arg Tyr Leu Leu His His Val Phe Glu Val Tyr Pro Tyr
100 105 110
       Val Glu Ser Gln Trp Ala Glu Ser Arg Gly Het Lys Tyr Lys Val Ser
55
       Thr Gly Leu Gln Ser Arg Tyr Arg Leu Val Asn Ser Asp Asn Cys Leu
130 135 140
       Het Phe Ala Thr Leu Gl; Val Phe Phe Glu Phe Glu Lys Trp Glu Gln 145 150 160
       Pro Ala Thr Ser Leu Phe Ala Gly Thr Tyr Ala Tyr Ser Arg Ser Ile
165 170 175
60
       Lys Ser His Leu Ser Ile Ser Phe Arg His Arg Leu Gly Glu His Trp
180 185 190
       Glu Phe Thr Thr Thr Ala Ile His Gln Gl; L;s Pro Asp Ser T;r Phe
195 200 205
65
        Lys Lys Ala Arg Phe Gly Gly Ala Ile Asp Leu Lys Tyr His Ile Thr
210 215 220
        Pro Thr Ile Gly Ile Arg Gly Ala Tyr Arg Ile Ile Tyr Asp Thr Ala
225 230 235 240
        Pro Ile Val Pro Val Arg Lys Asp Tyr Asn Thr Val Asp Val Gly Ile
245 255 255
 70
        Asp Ile Ser Phe
 75
        (2) INFORMATION FOR SEQ ID NO:489
```

```
(i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 834 amino acids
                    (B) TYPE: amino acid
  5
                    (D) TOPOLOGY: linear
             (ii) MOLECULE TYPE: protein
            (iii) HYPOTHETICAL: YES
 10
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...834
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 489
 20
        Het Lys Arg Ile Val Leu Ser Ser Phe Leu Phe Val Leu Ser Ile Leu
                                                  1.0
        Ser Leu Het Ala Gln Asn Asn Thr Leu Asp Val His Ile Ser Gly Thr 20 \hspace{1cm} 25 \hspace{1cm} 30
        Ile Lys Asp Ala Ser Ser Gly Glu Pro Val Pro Tyr Ala Thr Val Ser
35 40 45
 25
        Ile Arg Leu Thr Gly Ala Asp Thr Thr Gln Val Phe Arg Gln Val Thr 50 60
        Asp Gly Asn Gly Tyr Phe Val Ile Gly Leu Pro Ala Ala Pro Ser Tyr
65 70 75 80
30
        His Leu Thr Ala Ser Phe Val Gly Met Lys Thr His Thr Met Gln Ile
85 90 95
        Ser Arg Gly Asn Gly Gln His Asp Ile Lys Ser Ile Asp Ile Ser Leu
100 105 110
        Glu Ser Glu Asp Lys Gln Leu Ser Thr Val Thr Val Ser Ala Ala Arg
115 120 125
35
        Pro Leu Val Lys Met Glu Ile Asp Arg Leu Ser Tyr Asn Met Lys Asp
130 135 140
        Asp Pro Ala Ala Lys Thr Asn Asn Leu Leu Glu Met Leu Arg Asn Val
145 155 160
        Pro Leu Val Thr Val Asp Gly Gln Gly Asn Ile Gln Val Lys Gly Ser
165 170 175
40
        Ser Asn Phe Lys Ile His Leu Asn Gly Arg Pro Ser Thr Met Val Ser
180 185 190
       Ser Asn Fro Lys Glu Val Phe Arg Ser Ile Pro Ala His Thr Ile Lys
195 200 205
45
       Arg Val Glu Val Ile Thr Asp Pro Gly Val Lys Tyr Asp Ala Glu Gly 210 215 220
       Thr Ser Ala Ile Leu Asp Ile Val Thr Glu Glu Gly Lys Lys Leu Glu
225 230 236 240
                                                    235
       Gly Tyr Ser Gly Ser Ile Thr Ala Ser Val Ser Asn Asn Pro Thr Ala 245 250 255
50
       Asn Gly Ser Ile Phe Leu Thr Ala Lys Ser Gly Lys Val Gly Leu Thr 260 270
       Thr Asn Tyr Asn Tyr Tyr Gly Gly Lys Asn Lys Gly Ser Arg Tyr Phe
275 280 285
55
       Thr Glu Arg Thr Thr Ser Het Leu Gln Thr Ile Glu Glu Gly Lys Gly
290 295 300
       Gln Glu Thr Phe Gly Gly His Phe Gly Asn Ala Leu Leu Ser Phe Glu
305 310 315 320
60
       Ile Asp Ser Leu Asn Leu Phe Thr Val Gly Gly Asn Val Arg Leu Trp
325 330 335
       Glu Het Thr Thr Asp Arg Asn Ser Val Glu Lys Ser Phe Ala Gly Ser \frac{340}{100} \frac{345}{100} \frac{350}{100}
       Asn Leu Met Ser Tyr Ile Asp Arg Lys Leu Lys Thr Gln Met Asp Ala 355 360 365
65
       Gly Ser Tyr Glu Leu Asn Ala Asp Tyr Gln His Ser Thr Arg Leu Pro
370 375 380
       Gly Glu Leu Leu Thr Val Ser Tyr Arg Phe Thr His Asn Pro Asn Asn 385 390 395 400
70
       Ser Glu Thr Phe Ile Asp Gln Trp Lys Arg Asp Pro Leu Asn Thr Ala 405 410 415
       Asn Thr Ile Gln Tyr Ala Gly Gln His Ser Lys Ser Asp Ala Gly Met 420 425 430
       Asp Glu His Thr Ala Gln Val Asp Tyr Thr Arg Pro Leu Gly Gln Ala
75
                                        440
```

```
His Ser Leu Glu Ala Gly Leu Lys Tyr Ile Tyr Arg His Ala Thr Ser
450 455
                                  455
           450
       Asp Pro Leu Tyr Glu Ile Arg Pro Ser Glu Asp Ala Pro Trp Gln Pro
                            470
                                                    475
       465
       Gly Ser Leu Tyr Ala Gln Asn Pro Ser Asn Gly Lys Phe Arg His Asp
485 490 495
       Gln Tyr Ile Gly Ala Ala Tyr Ala Gly Tyr Asn Tyr Arg Lys Asp Gln
500 510
       Tyr Ser Leu Gln Thr Gly Leu Arg Val Glu Ser Ser Arg Leu Lys Ala
515 520 525
10
       Leu Phe Pro Glu Asn Ala Ala Ala Asp Phe Ser His Asn Ser Phe Asp
530 535 540
       Trp Val Pro Gln Leu Thr Leu Gly Tyr Thr Pro Ser Pro Met Lys Gln 545 550 560
       Leu Lys Leu Ala Tyr Asn Phe Arg Ile Gln Arg Pro Ala Ile Gly Gln
565 570 575
15
       Leu Asn Pro Tyr Arg Leu Gln Thr Asn Asp Tyr Gln Val Gln Tyr Gly
580 590
       Asn Pro Asp Leu Lys Ser Glu Lys Arg His His Val Gly Leu Ser Tyr
595 600 605
20
       Asn Gln Tyr Gly Ala Lys Val Net Leu Thr Ala Ser Leu Asp Tyr Asp
610 615 620
       Phe Cys Asn Asn Ala Ile Gln Asn Tyr Thr Phe Ser Asp Pro Ala Asn
625 630 635 640
       Pro Asn Leu Phe His Gln Thr Tyr Gly Asn Ile Gly Aro Glu His Ser
645 650 655
25
       Phe Ser Leu Asn Thr Tyr Ala Net Tyr Thr Pro Ala Val Trp Val Arg
660 665 670
       Ile Het Leu Asn Gly Asn Ile Asp Arg Thr Phe Gln Lys Ser Glu Ala
675 680 685
30
       Leu Gly Ile Asp Val Asn Ser Trp Ser Gly Het Val Tyr Ser Gly Leu 690 700

Het Phe Thr Leu Pro Lys Asp Trp Thr Val Asn Leu Phe Gly Gly Tyr 715 720
       Tyr His Gly Gly Arg Ser Tyr Gln Thr Lys Tyr Asp Gly Asn Val Phe 725 730 735
35
       Asn Asn Ile Gly Ile Ala Lys Gln Leu Phe Asp Lys Lys Leu Arg Val^{740} ^{745} ^{750}
       Ser Leu Ser Ala Asn Asn Ile His Ala Lys Tyr Ser Thr Trp Lys Ser
755 760 765
40
       Arg Thr Ile Gly Asn Gly Phe Thr Ile Tyr Ser Glu Asn Ala Gly Ile 770 780
       Gln Arg Ser Val Ser Leu Ser Leu Thr Tyr Ser Phe Gly Lys Met Asn
785 790 795 800
       Thr Gln Val Arg Lys Val Glu Arg Thr Ile Val Asn Asp Asp Leu Lys
805 810 815
45
       Gln Thr Ser Ser Gln Gly Gln Gln Gly Gly Gly Gln Gly Asn Pro Thr
       Gly Asn
50
        (2) IMFORMATION FOR SEQ ID NO:490
              (i) SEQUENCE CHARACTERISTICS:
55
                   (A) LENGTH: 399 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) HOLECULE TYPE: protein
60
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Perphyromonas gingivalis
65
             (i::) FEATURE:
                   (A) MANE/KEY: misc_feature
                   (B) LOCATION 1...399
70
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:490
       Net Arg Leu Ser Ala Ile Leu Ile Ala Leu Ile Val Met Leu Pro Ala
                                                 10
        Val Leu Ser Gly Gln His Tyr Tyr Ser Het Ala Gly Glu Arg Leu Glu
75
                                             25
```

```
Thr Asp Ser Ile Arg Pro Asn Glu Leu Ser Ala Ser Ile Arg Ser Ala 35 40 45
       Leu Phe Phe Arg Asn Asn Glu Tyr Asn Ala Arg Ser Val Lys Gly Tyr 50 60
       Thr Leu Pro Gly Ala Arg Val Ser Ala Phe Ala Ser Tyr Ser Leu Pro 65 70 80
       Ala Ala His Gly Val Lys Leu Ser Leu Gly Val Ser Thr Leu Asn Tyr
85 90 95
       Trp Gly Ala Ser Arg Tyr Pro Ala Gly Ile Ala Tyr Ser Asp Leu Pro 100 \hspace{1cm} 105 \hspace{1cm} 110
10
       Tyr Trp Thr Asp Tyr Asn Asp Tyr Val Arg Leu Arg Ile Leu Pro Tyr 115 120 125
       Val Gln Ala Net Leu Lys Pro Thr Ala Thr Thr Ala Leu Met Leu Gly
130 135 140
       Asn Ile Ala Gly Gly Thr Ala His Gly Leu Ile Glu Pro Ile Tyr Asn
145 150 150 160
15
       Pro Glu Leu Asp Leu Thr Ala Asp Pro Glu Ala Gly Val Gln Phe Arg
165 170 175
       Gly Asp Trp Thr Arg Phe Arg Met Asp Val Trp Val Asn Trp Net Ser
180 185 190
20
       Het Ile Phe Lys Asn Asp Asn His Gln Glu Ser Phe Val Phe Gly Leu
195 200 205
       Ser Thr Thr Ser Lys Leu Leu Ser Gly Glu Gly Lys Trp Arg Leu Glu 210 215 220
       Leu Pro Leu Gln Ala Ile Ala Thr His Arg Gly Gly Glu Tyr Asn Trp 225 230 235 240
       Ala Gln Gln Asp Thr Val His Thr Trp Val Asn Gly Ala Val Gly Leu 245 \hspace{1cm} 255 \hspace{1cm}
       Lys Leu Ser Tyr Arg Pro Arg Thr Asp Lys Pro Het Gln Ile Trp Gly 260 \phantom{000}265 \phantom{000}270
30
       Ser Ala Tyr Gly Val Ala Ala Leu Ser Ser Gly Gly Tyr Phe Pro Tyr 275 280
       Glu Arg Gly Trp Gly Gly Tyr Leu Ser Leu Gly Het Asp Leu Glu His
290 295 300
35
       Phe Ala Phe Arg Thr Asp Tyr Trp Tyr Gly Arg His Tyr Val Ser Pro 305 310 315
       Phe Ala Ala Pro Phe Ala Asn Ser Leu Thr Tyr Asp Lys Gln Pro Leu
325 330 335
       Thr Asn Gly Trp Gly Asp Tyr Ile Arg Leu Tyr Ala Asp Tyr Ser Trp 340 345 350
40
       Arg Het Ala Arg Ser Val Ser Leu Ala Ala Val Ala Arg Val Trp Phe 355 \hspace{1cm} 360 \hspace{1cm} 365
       Gln Pro Ser Asp Arg Phe Ala Net Ser His Ala Leu Glu Leu Thr Met
                                 375
                                                          380
45
       Arg Ile Asp Pro Lys Phe Pro Ile Ala Phe Leu Lys Gly Asn His
       (2) INFORMATION FOR SEQ ID NO:491
50
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 382 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
55
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
60
                   (A) ORGANISH: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (E) LOCATION 1...382
65
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:491
       Het Asn Lys Ser Leu Leu Ser Leu Ala Cys Leu Ile Leu Cys Gly Het 1 \phantom{\bigg|}5\phantom{\bigg|} 10 \phantom{\bigg|}15\phantom{\bigg|}
70
       Pro Ala Ile Ala Gln Gln Thr Gly Pro Ala Glu Arg Ser Gly Glu Pro
20 25 34
       Ser Leu Ala Glu Arg Val Phe Gly Leu Glu Gln Lys Gln Lys Leu
                                      4.0
       Lys Val Tyr Leu Gly Ile Gin Ser Phe Tyr Asp Gin Pro Leu Val Asp
75
```

```
Asp Glu Ser His Ile Gly His Phe Lys Val Gln Glu Leu Arg Het Ser 65 70 75 80
       Ala His Gly Glu Leu Asn Arg His Leu Ser Phe Asp Trp Arg Gln Arg
85 90 95
       Leu Asn Arg Ala Ala Asp Gly Thr Ser Phe Ala Asp Asn Leu Ser Asn
100 105 110
       Ala Ile Asp Ile Ala Gly Val Asp Trp His Pro Asn Asp Lys Val Ser
115 120 125
       Phe Phe Gly Arg Gln Tyr Ala Arg Phe Gly Gly Ile Glu Tyr Asp
130 135 140
10
      Net Asn Pro Val Glu Ile Tyr Gln Tyr Ser Asp Leu Val Asp Tyr Net
145 150 155 160
       Thr Cys Tyr Thr Ser Gly Val Asn Phe Ala Trp Asn Phe His Pro Glu
165 170 175
      Gln Gln Leu Gln Leu Gln Val Leu Asn Ala Tyr Asn Asn Arg Phe Ala 180

Asp Arg Tyr His Val Thr Pro Asp Val Ala Thr Ala Thr Ser Tyr Pro 200 205
15
       Leu Leu Tyr Ser Ala Gln Trp Asn Gly Thr Leu Leu Gly Gly Ala Leu
210 215 220
20
       His Met Arg Tyr Ala Val Ser Met Ala His Gln Ala Gln Glu Arg Asn
225 230 235 240
       Het Trp Tyr Phe Thr Ala Gly Asn Leu Phe Asn Pro Gly Lys Arg Ile
245 250 255
       Asn Gly Tyr Leu Asp Leu Thr Tyr Ser Ile Glu Gly Leu Asp Asp Lys
260 265 270
25
       Gly Ile Met Thr Ala Arg Tyr Gly Lys Gly Lys Thr Leu Thr Asp Val
275 280 285
       Lys Tyr Tyr Ala Leu Val Ser Lys Trp Asn Phe Arg Ile Phe Asp Gln 290 295 300
30
       Val Asn Leu Phe Leu Lys Gly Met Tyr Glu Asn Gly Tyr Ala Pro Ala
305 310 320
       Gln Tyr Gly Glu Ser Ser His Thr Arg His Ser Tyr Gly Tyr Het Gly
325 330 335
       Gly Val Glu Tyr Tyr Pro Thr Glu Thr Asn Phe Arg Leu Phe Val Thr 340 345 350
35
       Tyr Ile Gly Arg His Tyr Arg Tyr Ser Ala Thr Glu Thr Glu Ser Thr 355 360 365
       Asn Ala Leu Arg Ala Gly Leu Ile Tyr Gln Ile Pro Phe Leu
40
       (2) INFORMATION FOR SEQ ID NO:492
              (i) SEQUENCE CHARACTERISTICS:
45
                   (A) LENGTH: 222 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
50
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
55
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...222
60
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:492
       Het Lys His Leu Phe Lys Ser Thr Leu Val Leu Leu Cys Ala Leu Ser
       Phe Ser Gly Thr Tyr Thr Phe Ala Gln Glu Asn Asn Thr Glu Lys Ser
20 25 30
65
       Arg Phe Asp Phe Ser Val Arg Leu Gly Gin Gly Tyr Ile Ala Gly Ser 35 40 45
       Thr Thr Asn Leu Met Tyr Gly Tyr Thr Ser Ala Asn Asp Arg Leu Leu
50 60
       Ser Gly Ala Ile Tyr Leu Gly Leu Thr Pro Ser Lys Lys Glu Asn Ala
65 70 80
       Thr Gly Val Ala Phe Arg Phe Leu Ser Pro Ser Pro Gly Tyr Tyr Val
                         85
                                              90
       Asp Ile Ser Gly Lys Glu Asn Thr Leu Asn Tyr Ala Phe Tyr Val Val
 75
                                            105
```

```
Gly Ala Tyr Asn Arg Ile Ala Ile Fro Ile Arg Pro Ile Lys Asn Phe 115 125
      Asn Phe Ile Phe Ser Thr Glu Val Gly Net Ala Trp Net Ser Arg His
130 135 140
       Glu Gln Ile Tyr Asn Ser Thr Ser Gln Thr Trp Asp Lys Gln Arg Lys
145 150 155 160
 5
       Ser Arg Ser Gly Leu Asp Phe Gly Leu Gly Het His Leu Gln Unk His 165 170 175
       Ile Asn Lys Thr Val Tyr Phe Met Ala Gly Thr Asp Leu Thr Ser Cys 180 185 190
10
      Met Phe Gly Lys Arg Ile Asn Asp Tyr Gln Gln Lys Asp Arg Thr Phe 195 \hspace{1cm} 200 \hspace{1cm} 205 \hspace{1cm}
       Ile Ala Leu Ile Asp Asn Ser Ile Gly Ile Gly Leu Asn Leu
15
       (2) INFORMATION FOR SEQ ID NO:493
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 391 amino acids
20
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
25
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
30
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                  (B) LOCATION 1...391
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:493
35
      Leu Phe Leu Phe Ala Pro Lys Ala Met Lys Ala Gln Glu Gln Leu Asn
20 25 30
40
       Val Val His Thr Ser Val Pro Ser Leu Asn Ile Ser Pro Asp Ala Arg
35 40 45
       Ala Ala Gly Het Gly Asp Ile Gly Val Ala Thr Thr Pro Asp Ala Tyr
50 55 60
       Ser Gln Tyr Trp Asn Pro Ser Lys Tyr Ala Phe Met Asp Thr Lys Ala
65 70 75 80
45
       Gly Ile Ser Phe Ser Tyr Thr Pro Trp Leu Ser Lys Leu Val Asn Asp
85 90 95
       Ile Ala Leu Met Gln Met Thr Gly Phe Tyr Lys Leu Gly Thr Asp Glu
100 105 110
50
       Asn Gln Ala Ile Ser Ala Ser Leu Arg Tyr Phe Thr Leu Gly Lys Leu
115 120 125
       Glu Thr Phe Asp Glu Leu Gly Glu Ser Met Gly Glu Ala His Pro Asn
130 135 140
       Glu Phe Ala Val Asp Leu Gly Tyr Ser Arg Gln Leu Ser Glu Asn Phe
145 150 155 160
55
       145
       Ser Met Ala Val Ala Leu Arg Tyr Ile Arg Ser Asp Gln Ser Thr His 165 $170$
       Asn Thr Gly Glu Asn Gln Ala Gly Asn Ala Phe Ala Ala Asp Ile Ala
180 185 190
60
       Gly Tyr Leu Gln Lys Tyr Val Leu Leu Gly Asn Ala Glu Ser Leu Trp 195 200 205
       Ser Leu Gly Phe Asn Val Lys Asn Ile Gly Thr Lys Ile Ser Tyr Asp
210 215 220
       Gly Gly Val Thr Ser Phe Phe Ile Pro Thr Ser Leu Asn Leu Gly Thr
225 230 235 240
65
       Gly Leu Leu Tyr Pro Ile Asp Asp Tyr Asn Ser Ile Asn Phe Asn Leu 245 \hspace{1cm} 250 \hspace{1cm} 255
       Glu Leu Ser Lys Leu Leu Val Pro Thr Pro Pro Ile Net Asp Gln Asn 260 265 270
       Asp Gin Ala Gly Tyr Glu Ala Ala Leu Lys Lys Tyr Gln Glu Thr Ser
275 280 285
70
       Ser Ile Ser Gly Ile Phe Ser Ser Phe Gly Asp Ala Pro Gly Gly Leu
290 295 300
                               295
                                                   300
       Lys Glu Glu Phe Arg Glu Ile Thr Trp Gly Leu Gly Ala Glu Tyr Ser
75
```

```
Tyr Asp Asp Lys Phe Phe Val Arg Ala Gly Tyr Ser Tyr Leu His Pro
325 330 335
                         325
       Thr Lys Gly Asn Leu Gln Tyr Phe Thr Ala Gly Ala Gly Phe Lys Het
                    340
                                      345
       Asn Ile Phe Arg Ile Asp Ala Ser Tyr Leu Leu Ser Thr Ile Gln Ser
355 360 365
       Asn Pro Leu Asp Gln Thr Leu Arg Phe Thr Leu Ala Phe Asp Net Asp
                                375
       Gly Leu Arg Asn Leu Phe His
10
       (2) INFORMATION FOR SEQ ID NO:494
             (i) SEQUENCE CHARACTERISTICS:
15
                   (A) LENGTH: 446 amino acids
                  (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
20
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...446
30
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:494
       Met Lys Thr Thr Val Gln Gln Ile Ile Leu Cys Leu Ala Leu Met Het
       Ser Gly Val Leu Gly Gly Asn Ala Gln Ser Phe Trp Glu Glu Ile Ala
25 30
35
       Pro Pro Phe Ile Ser Asn Glu Pro Asn Val Lys Tyr Ile Ile Pro Asn 35 40 45
       Met Gly Ile Asp Ser Lys Gly Thr Ile Tyr Val Thr Val Thr Lys Arg
50 55 60
       Ile Gln Gln Gly Ala Asn Tyr Thr Ser Glu Gln Leu Gly Met Tyr Tyr
65 70 75 80
40
       Arg Pro Leu Gly Asp Asn Glu Gln Trp Trp Lys His Asp Pro Tyr Phe
85 90 95
       Asp Asp Lys Ile Val Ala Asp Ile Gln Thr Asp Ala Tyr Gly Arg Val
45
       Tyr Val Cys Thr Thr Ser Ser Arg Asp Gln Glu Tyr Gln Leu Tyr Ile
115 120 125
       Asn Glu Gln Asn Glu Trp Arg Cys Ile Phe Lys Thr Ser Val Ser Thr 130 135 140

Tyr Glu His Gly Met Ala Val Phe Arg Ser Ser Thr Gly Val Thr Tyr 145 150 150 155 160
50
       Ile Gly Thr Arg His His Ile Phe Ala Ser Gly Val Asn Asp Phe Glu
165 170 175
       Phe Asn Thr Ile Tyr Glu Asp Ser Thr Pro Net Ser Cys Arg Phe Ala
180 185 190
55
       Glu Ala Thr Asn Ser Gly Thr Ile Tyr Leu Ala Leu Het His Glu Thr
195 200 205
       Thr Net Ser Thr Thr Ile Leu Thr Tyr Gln Asn Gly Glu Phe Val Asp 210 215 220
60
       The Ser Glu Ser Glu Leu Ser Ash Ser The The Ala Ser Met Cys Ser 225 230 235 240
       Asn Lys Glu Giy Asp Ile Ile Ala Leu Val Thr Ser Tyr Thr Gly Fhe 245 250 255
       Het Ser Gly Thr Leu Ala Ile Arg Lys Ala Asp Glu Gly Lys Trp Gla
265 270
65
                     260
        Leu Val Gly Gly Asp Ile Gln Asn Ala Ile Val Gln Asn Ile Cys Met 275 -280 -285
       Het Asp Asp Asn Lys Ile Ala Cys Glu Val Phe Gly Thr Pro Asn Gl; 290 295 300
        Val Asp Gly Arg Thr Arg Val Cys Val Ser Asp Ala Ser Val Phe Asp
305 310 315 320
70
        Phe Glu Trp Tyr Glu Asp Glu Ile Tyr Gly Gly Leu Ile Phe Asp Thr 325 330 335
        Phe Phe Tyr Ser Pro Trp Asp Lys Leu Leu Tyr Ala Lys Phe Gly Gly
75
```

```
Ile Net Leu Arg Ser Lys Glu Ser Phe Ile Thr Ser Phe Ile Ser Pro
                                      360
                                                             365
        Thr Val Val Gln Gly Val Asp Val Tyr Thr Leu Ala Gly Lys Ile Arg
                                375
                                                     380
       Ile Glu Ser Glu Thr Pro Val Ser Glu Val Leu Leu Phe Asp Leu Ala
385 390 395 400
  5
        Gly Arg Net Val Leu Arg Gln Thr Ile Asp Asn Lys Ile Tyr Ser Asp 405 410 415
        Ile Asp Thr Ash Gly Leu Lys Arg Ser Gly Ile Tyr Val Val Ser Val
420 425 430
 10
        Arg Leu Ser Ser Gly Gln Val Phe Ser His Lys Val Gln Val
        (2) INFORMATION FOR SEQ ID NO:495
 15
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 308 amino acids
                   (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
 20
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
25
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
30
                  (B) LOCATION 1...308
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:495
       Het Ile Ile Arg Cys Leu Ile Arg Arg Pro Arg Thr Val Leu Phe Gly
35
       Leu Ile Phe Val Val Gly Leu Phe Ser Ala Net Ala Gln Glu Lys Lys
20 25 30
       Asp Ser Leu Ser Thr Val Gln Pro Val Pro Asn Ser Ser Met Val Glu
                                     40
                                                           45
40
       Gln Thr Pro Leu Leu Ser Ile Asp His Pro Val Leu Pro Ala Ser Phe
50 60
       Gln Asn Thr Arg Thr Leu Lys Arg Phe Arg Asp Lys His Leu Ser Asp
65 70 75 80
       Ala Leu Leu Asn Gly Leu Lys Pro His Arg Ser Ser Leu Gln Leu Asn 85 90 95
45
                                            90
       Glu Glu Leu Asn Phe Ala Ala Glu Arg Arg Asp Phe Val Ser Pro Leu
100 105 110
       Leu Gln Thr Arg His Ala Ala Gly Val Leu Ser Trp Arg Pro Thr Asp
115 120 125
50
       Arg Het His Phe Tyr Thr Ser Gly Asn Ile Gly Leu Gly His Asp Leu 130 135
       Leu Thr Gly Val Arg Lys Asp Phe Gly Trp Asn Ala Gly Ala Asp Phe
145 150 155
       Leu Leu Ser Gln Asn Leu Thr Ala His Val Gln Gly Gly Trp Gln Gln
165 170 175
55
       Asn Phe Gly Phe Ile Pro Met Thr Ala Val Asn Gly Gln Leu Arg Trp
180 185 190
       Gln Ala Thr Glu Arg Leu Ser Phe Thr Thr Gly Ile Asp Tyr Arg Gln
195 200 205
60
       Val Gln Trp Asn Ala Phe Asp Asn Arg Thr Phe Ser Leu Lys Gly Ser
210 215 220
       Ala Arg Tyr Glu Val Met Asp Asn Val Phe Val Asn Gly Phe Gly Ser
225 230 235 240
      Tyr Pro Leu Tyr Ser Ser Thr Arg Ser Gly Leu Asn Het Ala Val Pro 245 250 255
65
      Het His Gly Phe Gly Pro Gln Tyr Gly Gly Ser Leu Glu Leu Lys Val
      Ser Glu Arg Phe Gly Phe Ala Val Gly Net Glu Arg Glu Tyr Asn Ile
275 280 285
                                                          285
70
      Trp Thr Arg Arg Trp Glu Thr His Tyr Phe Ala Tyr Pro Val Phe Tyr
                                 295
      Gly Asp Lys Lys
75
       (2) INFORMATION FOR SEQ ID NO:496
```

```
(i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 315 amino acids
                  (B) TYPE: amino acid
 5
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
10
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
15
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...315
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 496
      Net Lys Thr Ash Arg Arg Tyr Ala Phe Val Leu Pro Leu Leu Leu Leu Leu 15
20
                                              10
      Thr Gly Leu Leu Ala Trp Gly Gln Asp Ser Ser His Gly Ser Asn Thr
      Ala Phe Ala Thr Asp Ser Ser Ser Arg Glu Leu Pro Thr Glu Gln Ser
35 40 45
25
      Ala Tyr Arg Ile His Ser Ala Tyr Met Val Gly Gly Gly Ser Ile
50 55
       Thr Arg Asp Thr Tyr Leu Ser Pro Leu Arg Tyr Gly Gly Trp Thr Leu 65 70 75 80
30
      As n Leu Leu Gly Glu Lys Thr Phe Pro Leu Lys Ala Ser Asp Ser Arg 85 90 95
      Trp Net Ile Arg Thr Gly His Glu Leu Asp Phe Ala Leu Met Asp Asn 100 \hspace{1cm} 105 \hspace{1cm} 110
       Pro Ala Asn Asn Ala His Phe Tyr Ser Leu Leu Tyr Asn Gly Ser Ala
115 120 125
35
       Ala Ala Leu Tyr Arg Leu Gly Ala Lys His Leu Arg Ala Ala Trp Het
130 140
       Asp Asn Leu Arg Leu Ala Phe Gly Pro Gly Leu Glu Ile Gly Leu Gl; 145 150 150 160
40
       Gly Ile Tyr Ser Thr Arg Asn Gly Asn Asn Pro Ala Thr Leu Lys Leu
165 170 175
       Tyr Thr Asn Ala Ile Ala Gln Ala Ser Ile Gly Tyr Tyr Val Pro Ser
180 185 190
       Glu Thr Phe Pro Leu Tyr Phe Arg Leu Leu Ser Gln Ile Asn Leu Phe
195 200 205
45
       Gly Ile Ala Tyr Gly Asn Gly Phe Gly Glu Ser Tyr Tyr Glu Asn Phe
210 215 220
       Leu Leu Asn Asn Gly Ile Ala Gly Ser Leu His Phe Thr Tyr Pro Gly 225 230 235
       Lys Phe Thr Arg Phe Thr Thr Leu Ile Thr Ala Asp Ile Pro Ile Arg 245 250 255
50
       Asn Phe Cys Thr Leu Arg Val Gly Tyr Arg Tyr Ser His Leu Gly Ser
260 265 270
       Ser Leu Asn Ala Leu Asp Thr Arg Ile His Ser His Thr Ala Phe Ile
275 280 285
55
       Gly Phe Val Thr Glu Phe Tyr Arg Phe Arg Gl; Arg Lys Ala Met Asn
290 295 300
       Thr Gly Arg Arg Thr Ser Leu Tyr Tyr His Asp
60
       (2) INFORMATION FOR SEQ ID NO:497
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 285 amino acids
65
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
70
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
75
            (ix) FEATURE:
```

```
(A) NAME/KEY: misc_feature
(B) LOCATION 1...285
```

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:497
 5
       Het Val Val Thr Leu Leu Val Ile Val Gly Ile Val Val Val Arg
                                              10
       Tyr Ser Leu Arg Val His Val His Lys Thr Gly Thr Val Val Ser Ala
20 25 30
       Ala Ile Phe Gly Phe Ile Leu Leu Gly Lys Thr Val Pro Cys Asp Thr 35 40 45
10
       Arg Asn Phe Phe Ser Ser Glu Ser Asp Glu Pro Glu Ser Arg Val Ala 50 55 60
       Thr Glu Ile Ala His Leu Cys Glu Ile Gly Phe Gln Ile His Ala Ser 65 70 75 80
15
       Ser Ile His Val Ala Val Arg Thr Asp Phe Gly Gln Ala Gly Ile His 85 90 95
       Cys Pro Met Ala Thr Asp Ala Ser Ala Thr Glu Phe Asp Arg Ser Ala
100 105 110
20
       Glu Cys Ala Glu Arg Thr Ser Ala Gln Ile Asp Thr Ala Ile Arg Ser
115 120 125
       Gln Ser Gln Ile Ile Arg Thr His Ile Asp Thr Cys Pro Lys Ser Ser 130 135 140

Gly Thr Ile Gly Gly Ser Thr His Thr Ser Leu His Leu Lys Val Phe 145 150 150 160
25
       Asp Gly Arg Gl; Glu Val Gly His Ile His Pro Lys Asp Gly Leu Arg
       Phe Gly Val Val Glu Gly Tyr Ser Ile Gly Ser Tyr Val Asp Ala Ile
180 185 190
30
       Gly lle Gly Ala Thr His Ala Lys Ala Gly Ile Ser Asp Thr Arg Thr 195 200 205
       Gly lle Ala Gly Gly Tyr His Gly Gly Ser Gln His Gln Gln Ile Gly 210 220
       Asp Val Thr Thr Ile Ile Gly Leu Gly Lys Phe Gly Leu Ala Asn Val
225 230 235 240
35
       Gly Val Ser Asp Arg Gly Phe Arg Arg Gly Ala Ser Ser Tyr Asp Leu
245 250 255
       Tyr Gly Leu Glu Leu His Ile Ala Lys Thr Ile Tyr Leu Ala Val Gly 260 265 270
40
       Asp Gly Cys Leu Cys Arg Gln Ala Glu Arg Gln Asp Gly
                275
                                      280
       (2) INFORMATION FOR SEQ ID NO:498
45
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 599 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
50
            (ii) HOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
55
                  (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...599
60
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:498
      65
       Glu Phe Glu Ala Gly Gly Val His Ala Ala Val Ser Arg Gln Glu Gln
20 25 30
       Glu Ala Ala Tyr Asp Lys Thr Leu Asn Thr Ile Lys Glu Lys Glu Val
                                     40
      Val Net Gly Arg Val Thr Ala Ile Asn Lys Arg Glu Val Val Ile Asn 50 55
70
      Val Gly Tyr Lys Ser Glu Gly Val Val Pro Ala Thr Glu Phe Arg Tyr
65 70 75 80
      Asn Pro Glu Leu Lys Val Gly Asp Glu Val Glu Val Tyr Ile Glu Asn
85 90
```

Gln Glu Asp Lys Lys Gly Gln Leu Val Leu Ser His Arg Lys Gly Arg

75

```
105
                      100
       Ala Ala Arg Ser Trp Glu Arg Val Asn Glu Ala Leu Glu Lys Asp Glu
115 120 125
       Ile Val Lys Gly Tyr Val Lys Cys Arg Thr Lys Gly Gly Met Ile Val
130 135 140
 5
       Asp Val Phe Gly Ile Glu Ala Phe Leu Pro Gly Ser Gln Ile Asp Val
145 150 155 160
       Arg Pro Ile Arg Asp Tyr Asp Ala Phe Val Glu Lys Thr Met Glu Phe 165 $170$
       10
       Lys Val Leu Île Glu Ala Glu Leu Glu Gln Gln Lys Lys Glu Île Île
195 200 205
       Gly Lys Leu Glu Lys Gly Gln Val Leu Glu Gly Ile Val Lys Asn Ile
210 215 220
15
       Thr Ser Tyr Gly Val Phe Ile Asp Leu Gly Gly Val Asp Gly Leu Ile
225 230 235 240
       His Ile Thr Asp Leu Ser Trp Gly Arg Val Ala His Pro Glu Glu Ile
245 250 255

Val Gln Leu Asp Gln Lys Ile Asn Val Val Ile Leu Asp Phe Asp Glu
260 265 270
20
       Asp Arg Lys Arg Ile Ala Leu Gly Leu Lys Gln Leu Met Pro His Pro 275 280 285
       Trp Asp Ala Leu Asp Ser Glu Leu Lys Val Gly Asp Lys Val Lys Gly 290 295 300

Lys Val Val Val Het Ala Asp Tyr Gly Ala Phe Val Glu He Ala Gln 305 310 320
25
       Gly Val Glu Gly Leu Ile His Val Ser Glu Het Ser Trp Thr Gln His
325 330 335
       Leu Arg Ser Ala Gln Asp Phe Leu His Val Gly Asp Glu Val Glu Ala 340 345 350
30
       Val Ile Leu Thr Leu Asp Arg Glu Glu Arg Lys Met Ser Leu Gly Leu
355 360 365
       Lys Gln Leu Lys Pro Asp Pro Trp Ala Asp Ile Glu Thr Arg Phe Pro 370 375 380
35
       Val Gly Ser Arg His His Ala Arg Val Arg Asn Phe Thr Asn Phe Gly 385 390 395 400
       Val Phe Val Glu Ile Glu Glu Gly Val Asp Gly Leu Ile His Ile Ser
405 415
       Asp Leu Ser Trp Thr Lys Lys Ile Lys His Pro Ser Glu Phe Thr Glu 420 425 430
40
       Val Gly Ala Asp Ile Glu Val Gln Val Ile Glu Ile Asp Lys Glu Asn
435 440 445
       Arg Arg Leu Ser Leu Gly His Lys Gln Leu Glu Glu Asn Pro Trp Asp
450 455 460
45
       Val Phe Glu Thr Val Phe Thr Val Gly Ser Ile His Glu Gly Thr Val 465 470 480
       Ile Glu Val Het Asp Lys Gly Ala Val Val Ser Leu Pro Tyr Gly Val
485 490 495
50
       Glu Gly Phe Ala Thr Pro Lys His Met Val Lys Glu Asp Gly Ser Gln 500 \, 505 \, 510
       Ala Val Leu Glu Glu Lys Leu Pro Phe Lys Val Ile Glu Phe Asn Lys
515 520 525
       Asp Ala Lys Arg Ile Ile Val Ser His Ser Arg Val Phe Glu Asp Glu
530 535 540
55
       Gln Lys Met Ala Gln Arg Glu Ala Asn Ala Glu Arg Lys Ala Glu Ala
545 550 560
       Lys Ala Ala Gln Lys Glu Ala Ala Ala Glu Ala Ala Asn Pro Ala Gln 565 570 575
60
       Ala Val Glu Lys Ala Thr Leu Gly Asp Leu Gly Glu Leu Ala Ala Leu
                   580
       Lys Glu Lys Leu Ser Glu Asn
65
       (2) INFORMATION FOR SEQ ID NO: 499
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 550 amino acids (B) TYPE: amino acid
70
                   (D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
75
```

```
(vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
 5
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...550
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:499
10
       Het Gly Arg Val Thr Ala Ile Asn Lys Arg Glu Val Val Ile Asn Val
                                                10
       Gly Tyr Lys Ser Glu Gly Val Val Pro Ala Thr Glu Phe Arg Tyr Asn
20 25 30
                  20
                                         25
       Pro Glu Leu Lys Val Gly Asp Glu Val Glu Val Tyr Ile Glu Asp Gln 35 40
15
       Glu Asp Lys Lys Gly Gln Leu Val Leu Ser His Arg Lys Gly Arg Ala
50 55 60
       Ala Arg Ser Trp Glu Arg Val Asn Glu Ala Leu Glu Lys Asp Glu Ile
65 75 80
20
       Val Lys Gly Tyr Val Lys Cys Arg Thr Lys Gly Gly Met Ile Val Asp
85 90 95
       Val Phe Gly Ile Glu Ala Phe Leu Pro Gly Ser Gln Ile Asp Val Arg
100 105 110
       Pro Ile Arg Asp Tyr Asp Ala Phe Val Glu Lys Thr Het Glu Phe Lys
115 120 125
25
       Ile Val Lys Ile Asn Gln Glu Tyr Lys Asn Val Val Ser His Lys
130 135 140
       Val Leu Ile Glu Ala Glu Leu Glu Gln Gln Lys Lys Glu Ile Ile Gly
145 150 155 160
30
       Lys Leu Glu Lys Gly Gln Val Leu Glu Gly Ile Val Lys Asn Ile Thr
165 170 175
       Ser Tyr Gly Val Phe Ile Asp Leu Gly Gly Val Asp Gly Leu Ile His
180 185 190
       Ile Thr Asp Leu Ser Trp Gly Arg Vai Ala His Pro Glu Glu Ile Val
35
       Gln Leu Asp Gln Lys Ile Asn Val Val Ile Leu Asp Phe Asp Glu Asp
210 215 220
       Arg Lys Arg Ile Ala Leu Gly Leu Lys Gln Leu Met Pro His Pro Trp 225 230 235 240
                                                 235
40
       Asp Ala Leu Asp Ser Glu Leu Lys Val Gly Asp Lys Val Lys Gly Lys
245 250 255
       Val Val Val Met Ala Asp Tyr Gly Ala Phe Val Glu Ile Ala Gln Gly 260 265 270
       Val Glu Gly Leu Ile His Val Ser Glu Met Ser Trp Thr Gln His Leu
275 280 285
45
       Arg Ser Ala Gln Asp Phe Leu His Val Gly Asp Glu Val Glu Ala Val 290 295 300
       Ile Leu Thr Leu Asp Arg Glu Glu Arg Lys Net Ser Leu Gly Leu Lys 305 310 315 320
50
       Gln Leu Lys Pro Asp Pro Trp Ala Asp 11e Glu Thr Arg Phe Pro Val
325 330 335
       Gly Ser Arg His His Ala Arg Val Arg Asn Phe Thr Asn Phe Gly Val 340 345 350
       Phe Val Glu Ile Glu Glu Gly Val Asp Gly Leu Ile His Ile Ser Asp
355 360 365
55
       Leu Ser Trp Thr Lys Lys Ile Lys His Pro Ser Glu Phe Thr Glu Val 370 375 380
       Gly Ala Asp Ile Glu Val Gln Val Ile Glu Ile Asp Lys Glu Asn Arg
385 390 395 400
60
       Arg Leu Ser Leu Gly His Lys Gln Leu Glu Glu Asn Pro Trp Asp Val
405 415
       65
       Gly Phe Ala Thr Pro Lys His Net Val Lys Glu Asp Gly Ser Gln Ala
450 455 460
                                455
                                                       460
       Val Leu Glu Glu Lys Leu Pro Phe Lys Val Ile Glu Phe Asn Lys Asp
465 470 475 480
70
       Ala Lys Arg Ile Ile Val Ser His Ser Arg Val Phe Glu Asp Glu Gln
485 490 495
       Lys Het Ala Gln Arg Glu Ala Asn Ala Glu Arg Lys Ala Glu Ala Lys 500 \hspace{1cm} 505 \hspace{1cm} 510
       Ala Ala Gln Lys Glu Ala Ala Ala Glu Ala Ala Asn Pro Ala Gln Ala
75
                                       520
                                                             525
```

```
Val Glu Lys Ala Thr Leu Gly Asp Leu Gly Glu Leu Ala Ala Leu Lys
            530
       Glu Lys Leu Ser Glu Asn
 5
        (2) INFORMATION FOR SEQ ID NO:500
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 458 amino acids
10
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
             (ii) HOLECULE TYPE: protein
15
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                     (A) ORGANISM: Porphyromonas gingivalis
20
             (ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...458
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:500
25
       Het Ile Val Asp Val Phe Gl; Ile Glu Ala Phe Leu Pro Gly Ser Gln
       lle Asp Val Arg Pro Ile Arg Asp Tyr Asp Ala Phe Val Glu Lys Thr
20 25 30
       Het Glu Phe Lys Ile Val Lys Ile Asn Glu Glu Tyr Lys Asn Val Val 35
30
       Val Ser His Lys Val Leu Ile Glu Ala Glu Leu Glu Gln Gln Lys Lys
50 55 60
       Glu Ile Ile Gly Lys Leu Glu Lys Gly Gl<br/>n Val Leu Glu Gly Ile Val 65 \phantom{000}70\phantom{000} 75 \phantom{0000}80\phantom{000}
35
       Lys Asn Ile Thr Ser Tyr Gly Val Phe Ile Asp Leu Gly Gly Val Asp 85 90 95
       Gly Leu Ile His Ile Thr Asp Leu Ser Trp Gly Arg Val Ala His Pro
100 105 110
       Glu Glu Ile Val Gln Leu Asp Gln Lys Ile Asn Val Val Ile Leu Asp 115 120 125

Phe Asp Glu Asp Arg Lys Arg Ile Ala Leu Gly Leu Lys Gln Leu Met 130 135 140
40

        Pro His Fro Trp Asp Ala Leu Asp Ser Glu Leu Lys Val Gly Asp Lys

        145
        150
        155
        160

        Val Lys Gly Lys Val Val Val Wet Ala Asp Tyr Gly Ala Phe Val Glu
        165
        170
        175

45
        Ile Ala Gln Gly Val Glu Gly Leu Ile His Val Ser Glu Met Ser Trp
180 185 190
50
        Thr Gln His Leu Arg Ser Ala Gln Asp Phe Leu His Val Gly Asp Glu
195 200 205
        Val Glu Aia Val Ile Leu Thr Leu Asp Arg Glu Glu Arg Lys Met Ser
210 215 220
        Leu Gly Leu Lys Gln Leu Lys Pro Asp Pro Trp Ala Asp Ile Glu Thr 225 230 235 249
55
       Arg Phe Pro Val Gl; Ser Arg His His Ala Arg Val Arg Asn Phe Thr
245 250 255
        Asn Phe Gly Val Phe Val Glu Ile Glu Glu Gly Val Asp Gly Leu Ile
260 265 270
60
       His Ile Ser Asp Leu Ser Trp Thr Lys Lys Ile Lys His Pro Ser Glu
275 280 285
        Phe Thr Glu Val Gly Ala Asp The Glu Val Gin Val The Glu Ile Asp
290 295 300
        Lys Glu Asn Arg Arg Leu Ser Leu Gly His Lys Gln Leu Glu Glu Asn 305 310 315
65
        Pro Trp Asp Val Phe Glu Thr Val Phe Thr Val Gly Ser Ile His Glu
325 335
        Gly Thr Val Ile Glu Val Het Asp Lys Gly Ala Val Val Ser Leu Pro 340 350
        Tyr Gly Val Glu Gly Phe Ala Thr Pro Lys His Het Val Lys Glu Asp
355 360 365
70
        Gly Ser Gln Ala Vai Leu Glu Glu Lys Leu Pro Phe Lys Val Ile Glu
370 375 380
        Phe Asn Lys Asp Ala Lys Arg Ile Ile Val Ser His Ser Arg Val Phe
385 390 395 400
75
```

WO 99/29870 PCT/AU98/01023

```
Glu Asp Glu Gln Lys Het Ala Gln Arg Glu Ala Asn Ala Glu Arg Lys
                        405
                                     410
       Ala Glu Ala Lys Ala Ala Gln Lys Glu Ala Ala Ala Glu Ala Ala Asn
                   420
                                        425
                                                              430
 5
       Pro Ala Gln Ala Val Glu Lys Ala Thr Leu Gly Asp Leu Gly Glu Leu
               435
                                   440
       Ala Ala Leu Lys Glu Lys Leu Ser Glu Asn
10
       (2) INFORMATION FOR SEQ ID NO:501
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 426 amino acids
                  (B) TYPE: amino acid
15
                  (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
20
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
                  (A) MAME/KEY: misc_feature
                  (B) LOCATION 1...426
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:501
30
       Het Glu Phe Lys Ile Val Lys Ile Asn Gln Glu Tyr Lys Asn Val Val
       Val Ser His Lys Val Leu Ile Glu Ala Glu Leu Glu Gln Gln Lys Lys
20 25 30
       Glu Ile Ile Gly Lys Leu Glu Lys Gly Gln Val Leu Glu Gly Ile Val
35 40
35
       Lys Asn Ile Thr Ser Tyr Gly Val Phe Ile Asp Leu Gly Gly Val Asp 50 55 60
       Gly Leu Ile His Ile Thr Asp Leu Ser Trp Gly Arg Val Ala His Pro
65 70 80
40
       Glu Glu Ile Val Gln Leu Asp Gln L;s Ile Asn Val Val Ile Leu Asp
85 90
       Phe Asp Glu Asp Arg Lys Arg Ile Ala Leu Gly Leu Lys Gln Leu Het
100 105 110
       Pro His Pro Trp Asp Ala Leu Asp Ser Glu Leu Lys Val Gly Asp Lys
115 120 125
45
                                  120
       Val Lys Gly Lys Val Val Val Met Ala Asp Tyr Gly Ala Phe Val Glu
130 135 140
       Ile Ala Gln Gly Val Glu Gly Leu Ile His Val Ser Glu Met Ser Trp 145 150 150 160
50
       Thr Gln His Leu Arg Ser Ala Gln Asp Phe Leu His Val Gly Asp Glu
165 170 175
       Val Glu Ala Val Ile Leu Thr Leu Asp Arg Glu Glu Arg Lys Met Ser
180 185 190
       Leu Gly Leu Lys Gln Leu Lys Pro Asp Pro Trp Ala Asp Ile Glu Thr
195 200 205
55
       Arg Phe Pro Val Gly Ser Arg His His Ala Arg Val Arg Asn Phe Thr 210 215 220
       Asn Phe Gly Val Phe Val Glu Ile Glu Glu Gly Val Asp Gly Leu Ile
225 230 235
60
       His Ile Ser Asp Leu Ser Trp Thr Lys Lys Ile Lys His Pro Ser Glu
245 250 255
       Phe Thr Glu Val Gly Ala Asp Ile Glu Val Gln Val Ile Glu Ile Asp
260 265 270
       Lys Glu Asn Arg Arg Leu Ser Leu Gly His Lys Gln Leu Glu Glu Asn
275 280 285
65
       Pro Trp Asp Val Phe Glu Thr Val Phe Thr Val Gly Ser Ile His Glu
290 295 300
       Gly Thr Val Ile Glu Val Net Asp Lys Gly Ala Val Val Ser Leu Pro
305 310 310
                                                  315
70
       Tyr Gl; Val Glu Gly Phe Ala Thr Fro Lys His Het Val Lys Glu Asp
325 330 335
                                         330
       Gly Ser Gln Ala Val Leu Glu Glu Lys Leu Pro Phe Lys Val Ile Glu
340 345
       Phe Asn Lys Asp Ala Lys Arg Ile Ile Val Ser His Ser Arg Val Phe 355 360 365
75
```

```
Glu Asp Glu Gln Lys Met Ala Gln Arg Glu Ala Asn Ala Glu Arg Lys
          370
                             37.5
                                                  380
      Ala Glu Ala Lys Ala Ala Gln Lys Glu Ala Ala Ala Glu Ala Ala Asn
                          390
                                             395
 5
      Pro Ala Gln Ala Val Glu Lys Ala Thr Leu Gly Asp Leu Gly Glu Leu
                    405
                                         410
      Ala Ala Leu Lys Glu Lys Leu Ser Glu Asn
10
      (2) INFORMATION FOR SEQ ID NO:502
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 240 amino acids
                 (B) TYPE: amino acid
15
                 (D) TOPOLOGY: linear
          (ii) MOLECULE TYPE: protein
         (iii) HYPOTHETICAL: YES
20
          (vi) ORIGINAL SOURCE:
                (A) ORGANISM: Porphyromonas gingivalis
          (ix) FEATURE:
25
                 (A) NAME/KEY: misc_feature
                 (B) LOCATION 1...240
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:502
30
      Met Lys Lys Ala Ile Leu Ser Gly Ala Ala Leu Leu Leu Gly Leu Cys
      Ala Asn Ala Gln Asn Val Gln Leu His Tyr Asp Phe Gly His Ser Ile
                 20
                                      25
      Tyr Asp Glu Leu Asp Gly Arg Pro Lys Leu Thr Thr Thr Val Glu Asn 35 40
35
      Phe Thr Pro Asp Lys Trp Gl; Ser Thr Phe Phe Phe Ile Asp Net Asp 50 55 60
      Tyr Thr Gly Lys Gly Ile Gln Ser Ala Tyr Trp Glu Ile Ser Arg Glu 65 70 75 80
40
      Leu Lys Phe Trp Gln Ala Pro Val Ser Ile His Leu Glu Tyr Asn Gly
85 90 95
      Gly Leu Ser Thr Ser Phe Thr Phe Gly His Asp Ala Leu Ile Gly Ala
                                      105
      Thr Tyr Thr Tyr Asn Asn Pro Ser Phe Thr Arg Gly Phe Thr Ile Thr
115 120 125
45
      Pro Het Tyr Lys His Leu Gl; Ala His Asp Phe His Thr Tyr Gln Ile
130 135 140
      Thr Gly Thr Trp Tyr Het His Fhe Leu Asp Gly Leu Leu Thr Phe Asn 145 150 150 160
      Gly Phe Leu Asp Leu Trp Gly Phe Pro Gln Glu Asn Pro Ile Gly Gly
165 170 175
50
      Pro Val Leu Lys Glu Gly Asp Lys Phe Val Phe Leu Ser Glu Pro Gln
180 185 190
      Phe Trp Ile Asn Leu Asn Arg Ile Lys Gly Ile Asp Lys Asp Phe Asn
195 200 205
55
                                   200
      Leu Ser Ile Gly Thr Glu Met Glu Ile Ser Arg Asn Phe Ala Arg Met
                             215
                                                    220
      Asp Lys Phe Ser Cys Ile Pro Thr Leu Ala Val Lys Trp Thr Phe Asn
60
       (2) INFORMATION FOR SEQ ID NO:503
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 434 amino acids (B) TYPE: amino acid
65
                 (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
70
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISH: Perphyromonas gingivalis
75
           (ix) FEATURE:
```

```
(A) MAME/KEY: misc_feature (B) LOCATION 1...434
```

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:503
 5
       Het Tyr Lys Asp Tyr Lys Gly Leu Tyr Ala Ser Leu Arg Trp Tyr Ala
1 10 15
                                                  10
       Leu Ile Ile Gly Leu Leu Phe Ala Ala Asp Gly Ile Gln Ala Gln Asn
20 25 30
       Asn Asn Phe Thr Glu Ser Pro Tyr Thr Arg Phe Gly Leu Gly Arg Leu 35 40 45
10
       Gly Glu Arg Thr Thr Ile Ser Gly His Ser Het Gly Gly Leu Gly Val
50 55
       Gly Leu Arg Gin Gly Thr Tyr Val Asn Ala Val Asn Pro Ala Ser Tyr
65 70 80
15
       Ser Ala Val Asp Ser Net Thr Phe Ile Phe Asp Phe Gly Ala Ser Thr
85 90 95
       Gly Ile Thr Trp Tyr Ala Glu Asn Gly Lys Lys Asp Asn Arg Lys Met
100 105 110
20
       Gly Asn Ile Glu Tyr Phe Ala Met Leu Phe Pro Ile Ser Lys Ser Ile
115 120 125
       Ala Met Ser Ala Gly Val Leu Pro Tyr Ser Ala Ser Gly Tyr Gln Phe
130 135 140
       Gly Ser Val Asp Gln Val Glu Gly Gly Ser Val Gln Tyr Thr Arg Lys
145 150 155 160
25
       Tyr Leu Gly Thr Gly Asn Leu Asn Asp Leu Tyr Val Gly Ile Gly Ala
165 170 175
       Thr Pro Fhe Lys Asn Phe Ser Ile Gly Ala Asn Ala Ser Ser Leu Phe
180 185 190
       Gly Arg Phe Thr His Ser Arg Gln Val Ile Phe Ser Thr Glu Ala Pro
195 200 205
30
       Tyr Asn Pro Val His Leu Ser Thr Leu Tyr Leu Lys Ala Ala Lys Phe
210 215 220
       Asp Phe Gly Met Gln Tyr His Leu Leu Leu Lys Ser Asp Arg Ser Leu
225 230 235 240
35
       Val Ile Gly Ala Val Tyr Ser Pro Arg Val Lys Met His Ser Glu Leu
245 250 255
       Thr Gln Ile Lys Asn Gln Val Gln Asn Gly Val Val Val Glu Ser Glu
260 265 270
       Thr Gln Glu Tyr Ile Lys Gly Met Asp Tyr Tyr Thr Leu Pro His Thr 275 280 285

Leu Gly Ile Gly Phe Ser Tyr Glu Lys Lys Asp Lys Leu Leu Gly 290 295 300
40
       Ala Asp Val Gln Tyr Ser Lys Trp Lys Gly Glu Lys Phe Tyr Lys Ser 305 310 315
45
       Asp Cys Lys Phe Gln Asp Arg Ile Arg Val Ser Leu Gly Gly Glu Ile
325 330 335
       Ile Pro Asp Ile Asn Ala Val Gly Net Trp Pro Lys Val Arg Tyr Arg 340 345 350
50
       Phe Gly Leu His Gly Glu Asn Ser Tyr Leu Lys Val Pro Thr Lys Gly 355 360 365
       Gly Val Tyr Gln Gly Tyr His Ile Val Gly Ala Val Phe Gly Ile Gly 370 375 380
       Ile Pro Leu Asn Asp Arg Arg Ser Phe Val Asn Val Ser Leu Glu Tyr
385 390 395 400
55
       Asp Arg Leu Ile Pro Lys Glu Gly Met Ile Lys Glu Asn Ala Leu Lys
405 415
       Leu Thr Phe Gly Leu Thr Phe Asn Glu Ser Trp Phe Lys Lys Leu Lys
60
```

(2) INFORMATION FOR SEQ ID NO:504

- 65 (i) SEQUENCE CHARACTERISTICS:

 (A) LENGTH: 926 amino acids

 (B) TYPE: amino acid

 (D) TOPOLOGY: linear
- 70 (ii) MOLECULE TYPE: protein
 - (iii) HYPOTHETICAL: YES
- 75 (vi) ORIGINAL SOURCE:
 (A) ORGANISH: Porphyromenas gingivalis

```
(ix) FEATURE:
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...926
 5
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:504
       Net Arg Ser Ile Tyr Gln Leu Leu Leu Ser Ile Leu Leu Ala Ser Leu
1 10 15
       Gly Phe Val Gly Leu Glu Ala Gln Gln Ala Gly Val Ala Gly Arg Val 20 25 30
10
       Leu Asp Glu Glu Gly Asn Pro Met Ile Gln Ala Asn Val Gln Leu Val
       Gln Ser Thr Gly Gln Val Ala Val Ala Ala Gly Ala Thr Asn Glu Lys
50 55 60
15
       Gly Leu Phe Ser Leu Lys Thr Ser Gln Glu Gly Asp Tyr Ile Leu Arg
65 70 80
       Val Ser Tyr Val Gly Tyr Thr Thr His Asp Glu Lys Ile Ser Leu Arg 85 \\ 90 \\ 95
       Asn Gly Gln Thr Ile Thr Leu Lys Asp Ile Ser Met Asn Glu Asp Ala
100 105 110
20
       Arg Leu Leu Gln Ser Val Thr Val Gln Ala Lys Ala Ala Glu Val Val
115 120 125
       Val Arg Asn Asp Thr Leu Glu Phe Asn Ala Gly Ser Tyr Thr Val Ala
130 135 140
25
       Gln Gly Ala Ser Ile Glu Glu Leu Ile Lys Lys Leu Pro Gly Ala Glu
145 150 155 160
       Ile Gly Ser Asp Gly Lys Ile Thr Ile Asn Gly Lys Asp Ile Ser Lys
165 170 175
       Ile Leu Val Asp Gly Lys Glu Phe Phe Ser Lys Asp Pro Gln Val Ala
180 185 190
30
       Ile Lys Asn Leu Pro Ala Asp Met Val Asn Lys Val Gln Val Leu Asn
195 200 205
       Lys Leu Ser Glu Leu Ser Arg Met Ser Gly Phe Asp Asp Gly Glu Glu 210 215 220
35
       Glu Thr Val Ile Asn Leu Thr Val Lys Pro Glu Lys Lys Gly Leu
225 230 235 240
        Phe Gly Thr Leu Gln Ala Gly Tyr Gly Thr Asp Gln Arg Tyr Het Ala
245 250 255
       Gly Gly Asn Val Asn Arg Phe Asp Gly Asn Lys Gln Trp Thr Leu Ile
260 265 270
40
        Gly Ser Ala Asn Asn Thr Asn Asn Met Gly Fhe Ser Glu Het Asp Ser
275 280 285
       Glu Het Gly Ser Met Thr Phe Phe Ser Pro Gln Gly Gly Gly Arg Arg
290 295 300

Gly Phe Gly Asn Ser Gly Gly Val Thr Ser Ser Ser Het Leu Gly Gly
305 310 315 320
45
        Asn Phe Ser Val Glu Phe Ser Ser Ala Leu Asn Thr Gly Gly Asp Ala
325 330 335
        Arg Tyr Gly Tyr Asn Asp Lys Ala Ile Glu Thr Thr Lys Arg Val Glu
340 350
50
        Asn Ile Leu Ala Glu Gly Asn Thr Tyr Met Asp Glu Asn Ile Leu Glu
355 360 365
        Arg Ser Phe Ser His Asn Gly Gln Ala Arg Phe Arg Met Gln Trp Lys
370 375 380
55
        Pro Ser Glu Arg Thr Glu Val Val Phe Glu Pro Asp Leu Ser Ile Ser
385 390 395 400
        Lys Ile Asp Gly Phe Phe Asn Asp Thr Tyr Glu Thr Lys Asp Ala Thr 405 410
        Gly Ile Ser Ile Asn Lys Gly Ser Ile His Gln Thr Thr Gln Gly Asn
420 425 430
60
        Asn Phe Arg Leu Asn Gly Glu Leu Asp Ile Ser His Lys Leu Asn Asp
445 445
        Glu Gly Arg Thr Ile Ser Ala Ser Val Ser Gly Gly Leu Thr Asp Glu
450 455 460
65
        Asp Gly Asp Gly Ile Tyr Gln Ala Val Leu Gln Ser Val Glu Thr Asn 465 470 475 489
        Gln Lys Gln Phe Asn Asp Asn Ser Asn Leu Gln Tyr Arg Leu Arg Leu
485 490 495
        Ser Tyr Val Glu Pro Leu Gly Lys Asn Tyr Phe Ala Gln Ala Ile Leu
500 505 510
70
        Asn Arg Arg Phe Ser Arg Arg Asn Ser Asp Arg Glu Val Tyr Arg Leu
515 520 525
        Gly Asp Asp Gly Gln Tyr Ser Ile Leu Asp Ser Gln Tyr Gly Leu Ser
 75
```

```
Tyr Ser Asn Glu Phe Thr Gln Tyr Arg Ile Gly Leu Asn Leu Lys Lys 545 550 555 560

Ile Ala Lys Thr Trp Asp Tyr Thr Val Gly Phe Asn Val Asp Pro Asn 575 575
       Arg Thr Val Ser Tyr Arg Ser Val Ala Gly Val Glu Gln Asp Lys Leu 580 585 590
       Ala Phe Asn Arg Val Asn Leu Ser Pro Het Leu Arg Ile Asn Tyr Lys
595 600 605
        Pro Ser Arg Thr Thr Asn Leu Arg Val Asp Tyr Arg Gly Arg Thr Thr 610 615 620
10
       Gln Pro Ser Ile Asn Gln Ile Ala Pro Val Gln Asp Ile Thr Asn Pro
625 630 635
       Leu Phe Val Thr Glu Gly Asn Pro Gly Leu Lys Pro Ser Tyr Ser Asn
645 650 655
15
       Asn Val Het Ala Het Phe Ser Asp Phe Asp Ala Lys Ser Gln Arg Ala 660 -665 -670
       Phe Asn Ile Val Phe Phe Gly Asn Tyr Thr Phe Asp Asp Ile Val Pro 675 680 685
       Asn Thr His Tyr Asp Pro Ser Thr Gly Ile Arg Thr Thr Arg Tyr Glu
690 695 700
20
       Asn Ala Ser Gly Thr Trp Gln Ala Asn Leu His Gly Thr Leu Ser Leu
705 710 715 720
       Pro Leu Lys Asn Arg Ala Phe Ser Phe Arg Het Ser Leu Phe Asn Arg 725 730 735
       Leu Ala Glu Gly Gln Ser Phe Ile Asn Asp Asp Lys Asn Lys Ala Leu 740 745 750
25
       Ser Fhe Arg Thr Arg Glu Arg Leu Thr Leu Thr Tyr Arg Asn Asn Trp
755 760 765
       lle Asp Thr Ser Ile Gly Gly Asn Ile Gly Phe Tyr Net Ala Asn Asn 776 -775
30
       Ser Leu Ser Gly Gln Lys Asp Ser Arg Thr Tyr Asp Phe Gly Gly Asn 785 790 795 800
       Tyr Gln Val Ala Leu Thr Leu Pro Tyr Gly Phe Arg Ile Asp Ser Asp 805 810 815
35
       Val Glu Tyr Asn Thr Asn Ser Gly Tyr Ser Gly Gly Phe Ser Leu Asp
820 825 830
       Glu Trp Leu Trp Asn Ala Ser Leu Ser Tyr Ser Phe Leu Arg Asp Lys
835 840 845
       Ala Gly Thr Leu Arg Val Asn Gly Tyr Asp Ile Leu Gly Gln Arg Ser 850 850
40
       Ser Ile Ser Arg Ser Ala Ser Ala Ile Asn Ile Glu Glu Ser Het Ser
865 870 875 880
       Asn Thr Ile Gly Arg Tyr Val Met Val Asp Phe Ile Tyr Arg Phe Asn 885
45
       Ala Phe Ser Gly Gly Gly Ser Arg Ser Asp His Gln Arg Gly Asn Het 900 \hspace{1cm} 905 \hspace{1cm} 910
       Asn Arg Fro Gly Pro Pro Phe Gly Gly Gly Arg Arg Pro Ser
915 920 925
50
       (2) INFORMATION FOR SEQ ID NO:505
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 400 amino acids
                    (B) TYPE: amino acid
55
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
60
            (vi) ORIGINAL SOURCE:
                   (A) ORGAHISH: Porphyromonas gingivalis
            (ix) FEATURE:
65
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...400
            (mi) SEQUENCE DESCRIPTION: SEQ ID NO:505
70
       Met Val Met Lys Leu Ile Lys Arg Ser Leu Leu Leu Leu Gly Ala Val
                                               10
       Leu Leu Ile Thr Leu Pro Ala Tyr Ser \overline{\text{Gin}} Asn Asp Asp 11e Phe Glu 25 \overline{\text{30}}
       Asp Asp Ile Tyr Thr Ser Arg Lys Glu Ile Arg Lys Gln Asn Gln Val
75
```

```
Lys Asp Trp Gln Asn Gln Glu Asp Gly Tyr Gly Asp Asp Thr Glu Tyr
50 60
       Thr Val Ala Ser Asp Arg Asp Ile Asp Ala Tyr Asn Arg Arg Asp Gly
65 70 75 80
       Gln Ser Tyr Asp Gly Lys Lys Leu Ser Lys Asp Lys Lys Arg Asp Ser 85\,
       Thr Arg Ser Ser Val Pro Gly Arg Tyr Ser Arg Arg Leu Ala Arg Phe
100 110
       Tyr Lys Pro Asn Thr Ile Val Ile Ser Gly Ala Asp Asn Val Tyr Val 115 120 125

Thr Asp Asp Gly Glu Tyr Phe Val Tyr Gly Asp Glu Tyr Tyr Asp Asp 130 135
10
       Ala Ser Ser Val Asn Ile Tyr Ile Asn Ser Pro Trp Cys Asp Pro Phe
145 150 155 160
       Pro Tyr Thr Ser Trp Tyr Pro Ser Fhe Ser Gly Trp Tyr Asn Tyr Thr
165 170 175
15
       Trp Asn Tyr Pro Trp Phe Tyr Tyr Gly Ser His Ile Gly Trp Gly Gly 180 185 190
                    180
       Tyr Tyr Pro Gly Tyr Asn Trp Tyr Trp Ser Tyr Tyr Tyr Asp Pro Phe 195 200 205
20
       Tyr Asn Pro Tyr Gly Ile Gly Met Gly Trp Gly Tyr Pro Tyr Gly Trp
210 215 220
       Gly Ser Tyr Tyr Gly Trp Gly Gly Tyr Pro Gly Val Ile His His Tyr
225 230 235
       His His Tyr Pro Lys Lys Thr Tyr Ser Asn Gly Gln His Ser Gly Ala
245 255
25
       Tyr Tyr Ser Tyr Gly Arg Pro Asn Arg Ile Lys Gly Gly Thr Ser Gly 260 265 270
       Ala Lys Leu Gly Thr Gly Arg Tyr Asp Arg Ile Gln Asn Ser Ser Ser
275 280 285
30
       Gln Lys Asn Lys Phe Gly Leu Gln Ser Asn Lys Pro Asn Asn Asn Leu 290 295 300
Gln Asn Val Lys Ser Gly Arg Thr Gly Arg Ala Asn Arg Asp Arg Asn 305 310 315 326
       Ile Glu Thr Val Thr Pro Asn Asn Gly Gln Lys Gln Asn Arg Pro Val
325 330 335
35
       Phe Gln Gln Asn Gln Ser Gly Asn Asp Arg Pro Thr Gly Arg Asn Ile
340 345 350
       Arg Ser Glu Arg Gln Gly Glu Asn Asn Asp Arg Thr Phe Ser Thr Pro
355 360 365
40
       Ser Arg Ser Asn Ser Asn Gly Gly Phe Ser Thr Pro Ser Arg Ser Ser
370 375 380
       Ser Gly Ser Het Ser Gly Gly Gly Gly Arg Ser Gly Arg Gly Arg Asm
45
       (2) INFORMATION FOR SEQ ID NO:506
              (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 398 amino acids
50
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromenas gingivalis
60
            (ix) FEATURE:
                   (A) MAME/KET: misc_feature
                   (B) LOCATION 1...398
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:506
65
       Met Lys Leu Ile Lys Arg Ser Leu Leu Leu Leu Gly Ala Val Leu Leu
       Ile Thr Leu Pro Ala Tyr Ser Gln Asn Asp Asp Ile Phe Glu Asp Asp 20 25 30
70
       The Tyr Thr Ser Arg Lys Glu Ile Arg Lys Gln Asn Gln Val Lys Asp 35 40 45
       Trp Gln Asn Gln Glu Asp Gly Tyr Gly Asp Asp Thr Glu Tyr Thr Val
50 55 60
       Ala Ser Asp Arg Asp Ile Asp Ala Tyr Asn Arg Arg Asp Gly Gln Ser
65 70 75
75
```

```
Tyr Asp Gly Lys Lys Leu Ser Lys Asp Lys Lys Arg Asp Ser Thr Arg
85 90 95
        Ser Ser Val Pro Gly Arg Tyr Ser Arg Arg Leu Ala Arg Phe Tyr Lys
100 105 110
        Pro Asn Thr Ile Val Ile Ser Gly Ala Asp Asn Val Tyr Val Thr Asp
115 120 125
 5
       Asp Gly Glu Tyr Phe Val Tyr Gly Asp Glu Tyr Tyr Asp Asp Ala Ser
130 135 140
       Ser Val Asn Ile Tyr Ile Asn Ser Pro Trp Cys Asp Pro Phe Pro Tyr 145 150 155 160
10
       Thr Ser Trp Tyr Pro Ser Phe Ser Gly Trp Tyr Asn Tyr Thr Trp Asn
165 170 175
        Tyr Pro Trp Phe Tyr Tyr Gly Ser His Ile Gly Trp Gly Gly Tyr Tyr
180 185 190

      Pro Gly Tyr Asn Trp Tyr Tyr Tyr Tyr Tyr Tyr Asp Pro Phe Tyr Asn 195
      200
      205

      Pro Tyr Gly Ile Gly Met Gly Trp Gly Tyr Pro Tyr Gly Trp Gly Ser 210
      215
      226

15
       Tyr Tyr Gly Trp Gly Gly Tyr Pro Gly Val Ile His His Tyr His His 225 230 235 240

Tyr Pro Lys Lys Thr Tyr Ser Asn Gly Gln His Ser Gly Ala Tyr Tyr 245 250 255
20
        Ser Tyr Gly Arg Pro Asn Arg Ile Lys Gly Gly Thr Ser Gly Ala Lys
260 265 270
25
       Leu Gly Thr Gly Arg Tyr Asp Arg Ile Gln Asn Ser Ser Ser Gln Lys 275 280 285
       Asn Lys Phe Gly Leu Gln Ser Asn Lys Pro Asn Asn Asn Leu Gln Asn
290 295 300
       Val Lys Ser Gly Arg Thr Gly Arg Ala Asn Arg Asp Arg Asn Ile Glu
305 310 315
30
       Thr Val Thr Pro Asn Asn Gly Gln L;s Gln Asn Arg Pro Val Phe Gln 325 330 335
       Gln Asn Gln Ser Gly Asn Asp Arg Pro Thr Gly Arg Asn Ile Arg Ser
340 345 350
35
       Glu Arg Gln Gly Glu Asn Asn Asp Arg Thr Phe Ser Thr Pro Ser Arg
355 360 365
        Ser Asn Ser Asn Gly Gly Phe Ser Thr Pro Ser Arg Ser Ser Ser Gly
370 375
        Ser Het Ser Gly Gly Gly Gly Arg Ser Gly Arg Gly Arg Asn
40
                                390
        (2) INFORMATION FOR SEQ ID NO:507
              (i) SEQUENCE CHARACTERISTICS:
45
                    (A) LENGTH: 581 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
50
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) NAME/KEY: misc feature
                    (B) LOCATION 1...581
60
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:507
       Het Ile Arg Lys Leu Ile Leu Leu Leu Ala Leu Het Pro Val Ala Ser
1 10 15
                                                   10
       Val Ala Phe Ala Val Pro Thr Asp Ser Thr Glu Ser Lys Asp Asn Arg
20 25 30
65
       Ile Leu Thr Ser Met Gln Ser Ser Ser Leu Asn Arg Asp Asp Ala Pro 35 40 45
                                        40
       Asp Lys Trp Gln Pro Het His Ala Asn Phe Ser Ile Gln Ser Asp Het 50 60
70
       Leu Leu Ser Thr Ala Gln Lys Ser Lys Asn Thr Trp Phe Gly Asn Ser
65 70 75 80
       Tyr Ile Met Gly Ile Ile Lys Asn Asn Tyr Leu Glu Phe Gly Ala Arg
85 90
        Phe Glu Asp Leu Tyr Lys Pro Leu Pro Gly His Glu Pro Glu Het Gly
75
                                               105
```

```
Arg Gly Val Pro His Met Tyr Val Lys Gly Ser Tyr His Trp Ala Glu
115 120 125
       Leu Thr Net Gly Asp Phe Tyr Asp Gln Phe Gly Ser Gly Net Val Phe
130 135 140
       Arg Thr Tyr Glu Glu Arg Asn Leu Gly Ile Asp Asn Ala Val Arg Gly 145 150 160
       Gly Arg Ile Val Leu Thr Pro Phe Asp Gly Val Arg Val Lys Gly Ile
165 170 175
       Ala Gly Gin Gin Arg Asn Tyr Phe Asp Arg Thr Gly Lys Val Phe Asn
180 185 190
10
       Ser Gly Arg Gly Tyr Leu Leu Gly Ser Asp Leu Glu Leu Asn Val Glu
195 200 205
       Arg Trp Ser Ser Ala Met Arg Asp Asn Asp Tyr His Leu Ala Ile Gly 210 215 220
       Gly Ser Phe Val Ser Lys His Glu Ala Asp Glu Asp Ile Phe Val Gly
225 230 235 240
15
       Val Gly Glu Asp Arg Lys Arg Leu Asn Leu Pro Leu Asn Val Pro Ile
245 250 255
       Het Gly Leu Arg Thr Asn Phe Gln Lys Gly Gly Leu Ala Leu Tyr Ala 260 265 270
20
       Glu Tyr Gly Tyr Lys Tyr Asn Asp Pro Ser Ala Asp Asn Asp Tyr Ile
275 280 285
       Tyr His Asp Gly Gln Ala Ala Leu Leu Ser Ala Ser Tyr Ser Lys Lys
290 295 300
25
       Gly Met Ser Ile Leu Leu Gln Ala Lys Arg Cys Glu Asn Phe Ala Phe 305 310 315
       Arg Ser Lys Arg Ser Ala Gln Leu Thr Pro Leu Met Ile Asn Tyr Met
325 330 335
       Pro Ala Phe Thr Gln Ala His Thr Tyr Thr Leu Ala Ala Ile Tyr Pro 340 345 350
30
       Tyr Ala Thr Gln Pro Gln Gly Glu Trp Ala Phe Gln Gly Glu Leu Arg 355 . 360 365
       Tyr Asn Phe Ala Arg Arg Thr Ala Leu Gly Gly Arg Tyr Gly Thr Gly 370 375 380
       Leu Arg Ile Asn Val Ser His Val Arg Gly Leu Asp Lys Lys Het Leu
385 390 395 400
35
       Lys Glu Asn Pro Asp Glu Leu Ile Gly Thr Asp Gly Tyr Thr Val Ser
405 415
       Phe Phe Gly Het Gly Asp Leu Tyr Tyr Ser Asp Ile Asp Val Glu Ile
420 425 430
40
       Thr Lys Lys Val Ser Pro Gly Phe Asn Phe Thr Leu Thr Tyr Leu Asn 435 440 445
       Gln Ile Tyr Asn Asn Lys Val Leu His Gly Ala Ala Gly Glu Lys Pro
450 455 460
       Glu Lys Ile Tyr Ala Asn Ile Phe Val Tyr Asp Gly Lys Tyr Lys Leu 465 470 475 480
45
       Ser Ash Lys Val Ala Leu Arg Thr Glu Leu Gln Tyr Leu His Thr Lys
485 490 495
       Gln Asp Gln Gly Asp Trp Ile Tyr Gly Met Ala Glu Leu Ser Ile Leu
500 510
50
       Pro Ser Leu Met Leu Ser Leu Ser Glu Gin Tyr Asn Ile Gly Glu Thr
515 520 525
       Lys Lys His Tyr Val Met Gly Ser Val Thr Tyr Thr His Gly Ala His 530 535 540
       Arg Val Ala Phe Ser Ala Gly Lys Thr Arg Ala Gly Met Asn Cys Ser 545 550 550 555 555 560 Gly Gly Val Cys Arg Val Val Pro Glu Thr Gln Gly Phe Tyr Leu Ser 565 565
55
       Tyr Ser Thr Asn Leu
60
        (2) INFORMATION FOR SEQ ID NO:508
              (i) SEQUENCE CHARACTERISTICS:
65
                    (A) LENGTH: 239 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) NOLECULE TYPE: protein
70
           (iii) HYFOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISM: Porphyromonas gingivalis
75
```

WO 99/29870 PCT/AU98/01023

460/490

```
(ix) FEATURE:
                 (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...239
 5
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:508
       Het Arg Ser Leu Phe Leu Ser Ala Leu Arg Ser Ser Ser Leu His Gly
                                             10
       Ser Glu Arg Arg Ser Arg Ile Ser Ser Ser Val Val Met Ser Ile Arg
10
                   20
                                       25
       Gln Lys Ile Arg Leu Phe His Leu Ser Val Cys Ala Gln Thr His Asp
             35
                                  40
       His Leu Ile Glu Ile His Leu Val Cys Ile Glu Phe Gly Ala Ile Asp
                            55
15
       Thr Asp Glu Phe Arg Leu Ser Ser His Ala Tyr Thr Thr Ser Pro Thr 65 70 80
       His Thr Gly Ala Ile His His Asn Cys Ile Glu Arg Ser Tyr Gly Arg
85 90
       Tyr Leu Val Thr Phe Gly Gln Glu Arg Asn Glu Leu His His Ser
20
                 100
                                       105
       Arg Pro Asp Arg Asn Ala Glu Val Tyr Arg Phe Pro Phe Asp Asn Ala
             115
                                   120
       The His Ser Ile Arg Tyr Glu Ala Phe Arg Pro Ile Arg Pro Ile Val
130 135 140
25
       Cys His Asp Asp His Phe Ile Ala Ile Gly Ser His Leu Phe Phe Lys
145 150 155 160
       Asp Asn Gln Ile Phe Ser Ser Gly Ser Gln Tyr Asp Asn Tyr Thr Val
165 170 175
       Ala Cys Phe Val Glu Ser Leu His Asp Arg Glu Gln Arg Ser His Thr
180 185
30
      His Thr Ala Ser Gly Thr Asn His Cys Ala Asp Leu Phe Asp Het Arg
195 200 205 205
Thr Leu Ser Gln Arg Thr Tyr His Ile Arg Asp Ile Val Ala Asp Phe
                              215
                                                     220
35
       Glu Phe Gly Gln Phe Leu Gly Arg Phe Ala His Gly Leu Asn His
                            230
       (2) INFORMATION FOR SEQ ID NO:509
40
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 211 amino acids
                 (B) TYPE: amino acid
                 (D) TOPOLOGY: linear
45
           (ii) NOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
50
                 (A) ORGANISM: Porphyromonas gingivalis
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                 (B) LOCATION 1...211
55
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:509
      Het Ser Ile Arg Gln Lys Ile Arg Leu Phe His Leu Ser Val Cys Ala
                                            10
60
      Gln Thr His Asp His Leu Ile Glu Ile His Leu Val Cys Ile Glu Phe
                                        25
      Gly Ala Ile Asp Thr Asp Glu Phe Arg Leu Ser Ser His Ala Tyr Thr
              35
                                   40
                                                         45
      Thr Ser Pro Thr His Thr Gly Ala Ile His His Asn Cys Ile Glu Arg
65
      Ser Tyr Gly Arg Tyr Leu Val Thr Phe Gly Gln Glu Arg Asn Glu Leu 65 70 75 80
       His His His Ser Arg Pro Asp Arg Asn Ala Glu Val Tyr Arg Phe Pro
85 90
                                            90
70
       Phe Asp Asn Ala Phe His Ser Ile Arg Tyr Glu Ala Phe Arg Pro Ile
100 105 110
       Arg Pro Ile Val Cys His Asp Asp His Phe Ile Ala Ile Gly Ser His
115 120
```

Leu Phe Phe Lys Asp Asn Gln Ile Phe Ser Ser Gly Ser Gln Tyr Asp 130 135 140

75

```
Asn Tyr Thr Val Ala Cys Phe Val Glu Ser Leu His Asp Arg Glu Gln
145 150 155 160
       Arg Ser His Thr His Thr Ala Ser Gly Thr Asn His Cys Ala Asp Leu
165 170 175
        Phe Asp Met Arg Thr Leu Ser Gln Arg Thr Tyr His Ile Arg Asp Ile
180 185 190
        Val Ala Asp Phe Glu Phe Gly Gln Phe Leu Gly Arg Phe Ala His Gly
10
        (2) INFORMATION FOR SEQ ID NO:510
               (i) SEQUENCE CHARACTERISTICS:
15
                      (A) LENGTH: 781 amino acids
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
             (ii) HOLECULE TYPE: protein
20
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                      (A) ORGANISM: Porphyromonas gingivalis
             (in) FEATURE:
                     (A) NAME/KEY: misc_feature
                      (B) LOCATION 1...781
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:510
30
        Het Asp Arg Pro Lys Pro Ser Tyr Ile Val Arg Ile Ala Ala Ile Leu
1 5 10 15
        Cys Leu Phe Val Gly Arg Pro Leu Phe Ala Gln Ser Tyr Val Asp Tyr
20 25 30
35
        Val Asp Pro Leu Ile Gly Thr Leu Ser Ser Phe Glu Leu Ser Ala Gly 35 40 45
        Asn Thr Tyr Pro Val Ile Gly Leu Pro Trp Gly Met Asn Ser Trp Thr 50 60
        Pro Met Thr Gly Val Pro Gly Asp Gly Trp Gln Tyr Thr Tyr Ser Ala
65 70 75 80
40
        His Lys Ile Arg Gly Phe Lys Gln Thr His Gln Pro Ser Pro Trp Ile
85 90 95
        85 90 95

Asn Asp Tyr Gly Gln Phe Ser Leu Leu Pro Leu Thr Ala Pro Gln Lys 100 105 110

Pro Ser Ser Asn Asp Ser IIe Ala Leu Thr Lys Trp Cys Lys Gln Leu 115 120 125

Phe Ser Asp Glu Gln Thr Ser Trp Phe Ser His Lys Ala Glu Thr Ala 130 140

Thr Pro Tyr Tyr Tyr Ser Val Tyr Leu Ala Asp Tyr Asp Thr Arg Val 145 150 150 155 160

Clu Het Ala Pro The Glu Arg Ala Ala IIe Phe Arg Ile Arg Tyr Ser
45
50
        Glu Het Ala Pro Thr Glu Arg Ala Ala Ile Phe Arg Ile Arg Tyr Ser
165 170 175
        Gly Asn Thr Glu Ser Gly Ser Gly Arg Trp Leu Arg Leu Asp Ala Phe
180 185
        Thr Gly Gly Ser Glu Ile Ser Ile Val Asp Pro His Thr Val Val Gly 195 200 205
        Ile Ser Arg Lys Asn Ser Gly Gly Val Pro Ala Asn Phe Ala Cys Tyr
210 215 220
        Phe Ile Leu Gln Ser Asp Thr Pro Net Ala Asp Val Leu Leu Glu Thr 225 230 230 240
60
        Asp Thr Gly Lys Ser Asp Glu Gly Thr Arg Ala Trp Ala Ala Cys Arg
245 255
        Phe Asp Ser Gln Glu Val Thr Val Arg Val Ala Ser Ser Phe Ile Ser 260 265 270
65
        Val Glu Gln Ala Glu Arg Asn Leu Ala Glu Val Lys Gly Gln Ser Phe
275 280 285
        Asp Arg Ile Arg Leu Ala Gly Arg Glu Ala Trp Asn Lys Val Leu Gly 290 295 300
        Arg Ile His Val Glu Gly Gly Thr Lys Asp Glu Arg Thr Thr Phe Tyr 305 310 315 320
70
        Ser Ala Leu Tyr Arg Cys Leu Leu Phe Pro Arg Arg Phe Tyr Glu Glu
325
330
335
Asp Ala Ser Gly Asn Phe Val His Tyr Ser Pro Tyr Asn Gly Glu Val
340
345
350
75
```

```
Leu Pro Gly Tyr Leu Tyr Thr Asp Thr Gly Phe Trp Asp Thr Phe Arg 355 360 365
       Ala Leu Phe Pro Leu Leu Asn Leu Leu Tyr Pro Asp Glu Asn Ile Lys
370 375 380
       Ile Gln Glu Gly Leu Asn Val Tyr Arg Glu Ser Gly Phe Phe Pro 385 390 395 400
       Glu Trp Ala Ser Pro Gly His Arg Asp Cys Met Ile Gly Asn Asn Ser
405 410 415
       Ala Ser Val Leu Ala Asp Ala Tyr Leu Lys Gly Val Arg Val Glu Asp
420 425 430
10
       Thr Arg Thr Leu Met Asn Gly Leu Leu His Ala Thr Lys Ala Val His
435 440 445
       Pro Lys Ile Ser Ser Thr Gly Arg Lys Gly Trp Glu Trp Tyr Asn Ser
450 455 460
15
       Leu Gly Tyr Val Pro Ala Asp Ala Gly Ile Asp Glu Ser Ala Ala Arg 465 $470$ 470 480
       Thr Leu Glu Tyr Ala Tyr Asn Asp Trp Cys Ile Leu Arg Leu Gly Arg
485 490 495
       Thr Leu Gly Trp Asp Arg Ala Ala Leu Asp Thr Leu Ala His Arg Ser
500 505 510
20
       Het Asn Tyr Arg His Leu Phe Asp Pro Glu Thr Lys Leu Met Arg Gly 515 520 525
       Arg Asn Gln Asp Gly Ser Phe Arg Thr Pro Phe Ser Pro Phe Lys Trp
530 535 540
       Gly Asp Val Phe Thr Glu Gly Asn Ala Trp His Tyr Thr Trp Ser Val
545 550 560
25
       Phe His Asp Val Gln Gly Leu Ile Asp Leu Het Gly Gly Asp Arg Pro
575 576
       Phe Val Ser Met Leu Asp Ser Val Phe Asn Thr Pro Pro Met Phe Asp 580 585 590
30
       Glu Ser Tyr Tyr Gly Phe Val Ile His Glu Ile Arg Glu Met Gln Ile
595 600 605
       Ala Asp Het Gly Asn Tyr Ala His Gly Asn Gln Pro Ile Gln His Met 610 615 620
35
       Ile Tyr Leu Tyr Asn His Ala Gly His Pro Trp Lys Ala Gln Glu Arg 625 630 635 640
       Leu Arg Glu Val Met Gly Arg Leu Tyr Arg Pro Thr Pro Asp Gly Tyr
645 650 655
       Cys Gly Asp Glu Asp Asn Gly Gln Thr Ser Ala Trp Tyr Val Phe Ser
40
       Ala Leu Gly Phe Tyr Pro Val Thr Pro Ala Thr Asp Gln Tyr Val Leu
675 680 685
       Gly Ser Pro Ile Phe Ser Lys Val Ile Leu Ser Phe Pro Asp Gly His
690 695 700
       Lys Thr Val Leu His Ala Pro Ala Asn Ser Ala Asp Thr Pro Tyr Ile
705 710 720
45
       Arg Ser Ile Ser Val Glu Gly Lys Glu Trp Ser Cys Asn Tyr Leu Thr
725 730 735
       His Glu Gln Leu Arg Ser Ser Ala Ser Ile Gln Trp Met Het Asp Thr
740 745 750
50
       Lys Pro Asn Tyr Asn Arg Gly Met Lys Glu Ser Asp Arg Pro Tyr Ser
785 760 765
       Phe Ser Thr Glu Gln Gln Arg Arg Ala Asn His Ser Asn
55
       (2) INFORMATION FOR SEQ ID NO:511
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 271 amino acids
60
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
65
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
70
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...271
            (:i) SEQUENCE DESCRIPTION: SEQ ID NO:511
75
```

```
Met Het Lys Ser Met Arg Ser Val Leu Leu Leu Phe Pro Leu Ser
1 5 10 15
      Leu Ile Thr Ala Leu Gly Cys Ser Asn Asn Lys Ala Ala Glu Ser Lys
20 25 30
      Ser Val Ser Phe Asp Ser Ala Tyr Leu Glu Arg Tyr Ile Pro Leu Arg
35 40
      Ala Asp Ile Asp Thr Pro Ser Leu His Val Net Ile Ser Tyr Val Tyr
50 55 60
      Pro Ser Gly Asp Met Leu Thr Glu Ile Phe Asn Gly Leu Leu Phe 65 70 80
10
      Gly Asp Ser Leu Met Asp Ser Ser Ser Pro Glu Asn Ala Met Glu Gly
85 90 95
      Tyr Ala Gln Met Leu Gly Glu Asp Tyr Arg Ser Asn Asn Ala Glu Ala
100 105 110
      Asn Leu Gln Gly Leu Pro Ser Asp Leu Leu Asp Tyr Ile Tyr Lys Gln
115 120 125
15
      Glu Asn Thr Ile Ala Tyr Cys Asp Thr Gly Leu Ile Ser Thr Arg Ile
130 135 140
      Asn Thr Tyr Thr Tyr Glu Gly Gly Ala His Thr Glu Asn Thr Val Arg
145 150 150 155
20
      Phe Ala Asn Ile Leu Arg Thr Thr Gly Lys Val Leu Glu Glu Arg Asp
165 170 175
       Ile Phe Lys Ile Asp Tyr Ala Glu Arg Leu Ser Ala Leu Ile Ile Gly
180 185 190
      Gln Leu Val His Asp Phe Gly Lys Thr Thr Pro Ala Glu Leu Asp Ala
195 200 205
25
       Ile Gly Phe Phe Asn Ala Glu Glu Ile Gln Pro Asn Gly Asn Phe Het 210 215 220
      Ile Asp Asp Lys Gly Leu Thr Tyr Cys Phe Asn Glu Tyr Gln Ile Ala
225 230 235 240
30
      Ala Tyr Ala Arg Gly Ala Val Tyr Val Arg Leu Gly Tyr Asp Val Leu
245 250 255
      Ala Pro Leu Leu Arg Asp Asp Ser Pro Leu Lys Arg Tyr Leu Pro
35
       (2) INFORMATION FOR SEQ ID NO:512
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 270 amino acids
40
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGAHISH: Porphyromonas gingivalis
            (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...270
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:512
55
       Net Lys Ser Met Arg Ser Val Leu Leu Leu Leu Phe Pro Leu Ser Leu l1 \\ 5 \\ 10 \\ 15
       Ile Thr Ala Leu Gly Cys Ser Asn Asn Lys Ala Ala Glu Ser Lys Ser 20 \\ 25 \\ 30
60
       Val Ser Phe Asp Ser Ala Tyr Leu Glu Arg Tyr Ile Pro Leu Arg Ala
35 40 45
       Asp Ile Asp Thr Pro Ser Leu His Val Het Ile Ser Tyr Val Tyr Pro 50 60
       Ser Gly Asp Asp Met Leu Thr Glu Ile Phe Asn Gly Leu Leu Phe Gly 65 70 75 80
65
       Asp Ser Leu Net Asp Ser Ser Ser Pro Glu Asn Ala Met Glu Gly Tyr
85 90 95
       Ala Gl<br/>n Met Leu Gly Glu Asp Tyr Arg Ser Asn Asn Ala Glu Ala As<br/>n 100 105 110
70
       Leu Gln Gly Leu Pro Ser Asp Leu Leu Asp Tyr Ile Tyr Lys Gln Glu
115 120 125
       Asn Thr Ile Ala Tyr Cys Asp Thr Gly Leu Ile Ser Thr Arg Ile Asn
130 135 140
       Thr Tyr Thr Tyr Glu Gly Gly Ala His Thr Glu Asn Thr Val Arg Phe
75
```

WO 99/29870 PCT/AU98/01023

```
Ala Asn Ile Leu Arg Thr Thr Gly Lys Val Leu Glu Glu Arg Asp Ile
165 170 175
       Phe Lys Ile Asp Tyr Ala Glu Arg Leu Ser Ala Leu Ile Ile Gly Gln
180 180 190
       Leu Val His Asp Phe Gly Lys Thr Thr Pro Ala Glu Leu Asp Ala Ile
195 200 205
 5
       Gly Phe Phe Asn Ala Glu Glu Ile Gln Pro Asn Gly Asn Phe Het Ile
210 215 220
       Asp Asp Lys Gly Leu Thr Tyr Cys Phe Ash Glu Tyr Gln Ile Ala Ala
225 230 235 240
10
       Tyr Ala Arg Gly Ala Val Tyr Val Arg Leu Gly Tyr Asp Val Leu Ala
245 250 255
       Pro Leu Leu Arg Asp Asp Ser Pro Leu Lys Arg Tyr Leu Pro
15
       (2) INFORMATION FOR SEQ ID NO:513
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 267 amino acids
20
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
25
          (jii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
30
            (ix) FEATURE:
                  (A) HAHE/KEY: misc feature
                  (B) LOCATION 1...267
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:513
35
       Het Arg Ser Val Leu Leu Leu Phe Pro Leu Ser Leu Ile Thr Ala
                                                10
       Leu Gly Cys Ser Asn Asn Lys Ala Ala Glu Ser Lys Ser Val Ser Phe 20 25 30
40
       Asp Ser Ala Tyr Leu Glu Arg Tyr Ile Pro Leu Arg Ala Asp Ile Asp 35 40 45
       Thr Pro Ser Leu His Val Met Ile Ser Tyr Val Tyr Pro Ser Gly Asp
50 55 60
       Asp Met Leu Thr Glu Ile Phe Asr Gly Leu Leu Phe Gly Asp Ser Leu 65 70 75 80
45
       Net Asp Ser Ser Ser Pro Glu Asn Ala Het Glu Gly Tyr Ala Gln Met
85 90 95
       Leu Gly Glu Asp Tyr Arg Ser Asn Asn Ala Glu Ala Asn Leu Gln Gly
100 105 110
       Leu Pro Ser Asp Leu Leu Asp Tyr Ile Tyr Lys Gln Glu Asn Thr Ile
115 120 125
50
       Ala Tyr Cys Asp Thr Gly Leu Ile Ser Thr Arg Ile Asn Thr Tyr Thr
130 135 140
       T_{Y}r Glu Gl_Y Gly Ala His Thr Glu Asn Thr Val Arg Phe Ala Asn Ile 145 150 150
55
       Leu Arg Thr Thr Gly Lys Val Leu Glu Glu Arg Asp Ile Phe Lys lle
165 170 175
       Asp Tyr Ala Glu Arg Leu Ser Ala Leu Île Île Gly Gln Leu Val His
180 185 190
       Asp Phe Gly Lys Thr Thr Pro Ala Glu Leu Asp Ala Ile Gly Phe Phe
195 200 205
60
       Asn Ala Glu Glu Ile Gln Pro Asn Gly Asn Phe Het Ile Asp Asp Lys 210 215 220
       Gly Leu Thr Tyr Cys Phe Asn Glu Tyr Gln Ile Ala Ala Tyr Ala Arg
225 230 235 240
65
       Gly Ala Val Tyr Val Arg Leu Gly Tyr Asp Val Leu Ala Pro Leu Leu
245 250 255
       Arg Asp Asp Ser Pro Leu Lys Arg Tyr Leu Pro
260 265
70
       (2) INFORMATION FOR SEQ ID NO:514
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 259 amino acids
75
                  (B) TYPE: amino acid
```

```
(D) TOFOLOGY: linear
           (ii) NOLECULE TYPE: protein
 5
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
10
           (ix) FEATURE:
                  (A) MAME/KEY: misc feature
                  (B) LOCATION 1...259
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:514
15
      Met Lys Lys Thr Thr Leu Thr Gly Ser Ile Cys Ala Leu Leu Leu Phe
1 5 19
       Leu Gly Leu Ser Ala Asn Ala Gln Ser Lys Leu Lys Ile Lys Ser Ile 20 $25$
20
       Glu Ala Ala Thr Thr Phe Ser Ser Ala Thr Ala Gly Asn Gly Fhe Gly
                                     40
       Gly Asn Ile Phe Gly Met Asp Met Ser Ile Arg Met Arg Val His His 50 60
       Ser Ile Leu Pro Glu Gly Leu Asp Phe Ser Val Gly Ile His Glu Arg
65 70 75 80
25
      Arg Ala His Trp Glu Glu Ala Gly Ser Pro Lys Leu Het Tyr Thr Asn
85 90 95
       Val Pro Ser Ile Ile Gly Ile Val Glu Lys Val Ile Val Phe Glu Asp
100 105 110
30
       Ala Glu Asp Phe Phe Asp Lys Lys Ala Leu Gly Arg Phe Leu Ile Ser 115 120 125
       Leu Gly Ile Ser Tyr Thr Lys His Leu Gly Ala Tyr Trp Gly Trp Thr
130 135 140
       Asn Asp Ala His Ile Leu Phe Ser Pro Ile Pro Lys Ser Lys Val His
145 150 155 160
35
       Tyr Asp Thr Tyr Thr Arg Ala Gly Ser Asp Leu Val Leu Gln Ser Glu
165 170 175
       Asp Val Ala Thr Val Ser Asn Gly Phe Ser Pro Gly Ile Gly Leu Lys
180 185 190
       Ser Ser Ile Trp Trp Lys Net Pro Ile Lys Ser Lys Tyr Asp Phe Arg
195 200 205
40
       Leu Gly Phe Ser Leu Gly Tyr Glu Tyr Leu Asn Leu Leu Tyr Pro Tyr
210 215 220
       Arg Asn Phe Lys Leu Asp Gly Asn Lys Pro Leu Ser Ala Leu Ser Pro 225 230 235 240
45
       Arg Net Asn His Ile Gly His Val Gly Phe Asn Phe Thr Val Gly Leu
       Trp Thr Asn
50
       (2) INFORMATION FOR SEQ ID NO:515
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 1266 amino acids
55
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) HOLECULE TYPE: protein
60
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
65
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...1266
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:515
70
       Met Gly Lys Tyr Lys Arg Ala Lys Tyr Arg Tyr Trp Leu Phe Pro Phe 1 \phantom{-}5\phantom{+} 10 \phantom{-}15\phantom{+}
       Cys Ser Asp Tyr Tyr Thr Phe Glu Gly Val Thr Phe Leu Cys Ala Ser 20 25 30
75
       Asp Asp Net Thr Thr Lys Lys Pro Gln Ala Ile Leu Asp Leu Glu Lys
```

WO 99/29870 PCT/AU98/01023

| | | | 2.5 | | | | | | | | | | | | | |
|----|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------|------------|------------|
| | Ala | Tyr 50 | 35 Asn | Ile | Glu | Ile | Pro 55 | 40 Asp | Leu | Ser | Ser | Gln 60 | 45 Glu | Gly | Ile | Ser |
| 5 | Trp 65 | | Val | Asn | Arg | Tyr 70 | | Lys | Gln | Asp | Ser 75 | | GJ 7. | Ala | Val | Val 80 |
| | Glu | Leu | Cla | Leu | Arg 85 | | Cys | Gln | Ile | Gl u 90 | | Het | Thr | Trp | Leu 95 | |
| 10 | | | | 100 | | | | | 105 | | | _ | | 110 | Ile | |
| 10 | | | 115 | | | | | 120 | | | | | 125 | | Arg | |
| | | 130 | | | | | 135 | | | | | 140 | | | Thr | |
| 15 | 145 | | | | | 150 | | | | | 155 | | | | Glu | 160 |
| | | | | | 165 | | | | | 170 | | | | | λεn 175 | |
| 20 | | | | 180 | | | | | 185 | | | | | 190 | Thr | |
| 20 | | | 195 | | | | | 200 | | | | | 205 | | Arg | |
| | | 210 | | | | | 215 | | | | | 220 | | | Lys | |
| 25 | 225 | | | | | 230 | | | | | 235 | | | | Arg | 240 |
| | | | | | 245 | | | | | 250 | _ | | | | Leu 255 | |
| 30 | | | | 260 | | | | | 265 | | | | | 270 | Leu | |
| 00 | | | 275 | | | | | 280 | | | | | 285 | | Ile | |
| | | 290 | | | | | 295 | | | | | 300 | | | Arg Thr | |
| 35 | 305 | | | | | 310 | | | | | 315 | | | | Glu | 320 |
| | | | | | 325 | | | | | 330 | | | | | 335 Asn | |
| 40 | | | | 340 | | | | | 345 | | | | | 350 | Lys | |
| | | | 355 | | | | | 360 | | | | | 365 | | Ser | |
| | | 370 | | | | | 375 | | | | | 380 | | | Lys | |
| 45 | 385 | | | | | 390 | | | | | 395 | | | | Leu | 400 |
| | | | | | 405 | | | | | 410 | | | | | 415 Leu | |
| 50 | | | | 420 | | | | | 425 | | | | | 430 | Leu | |
| | | | 435 | | | | | 440 | | | | | 445 | | Ile | |
| | | 450 | | | | | 455 | | | | | 460 | | | Thr | |
| 55 | 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| | | | | | 485 | | | | | 490 | | | | | 495 Leu | |
| 60 | Glu | Lys | Glu | 500 Lys | Glu | Lys | Gln | Lys | 505 Lyε | Thr | Ser | Val | Glu | 510 T;r | His | Pro |
| | Phe | Cys | 515 L;:s | Val | Met | Leu | | 520 Gly | Asn | His | Ser | | 525 Gly | Lys | Thr | Thr |
| 65 | Phe | 530 Leu | Ser | Gln | туr | Asp | 535 Thr | Asn | Туг | Thr | | 540 Gln | Lys | A <i>s</i> n | Thr | His |
| 03 | 545 Vāl | Leu | Ser | lle | His | 550 Arg | Ser | Asn | Asn | | 555 Asn | Alā | Ile | Phe | Tyr | 560 Asp |
| | Phe | Gly | Glγ | Gln | 565 Asp | Tyr | Tyr | His | | 570 Ile | Tyr | Gln | Ala | | 575 Phe | Thr |
| 70 | Thr | Gln | Ser 595 | 580 Leu | Tyr | Leu | Leu | | 585 Trp | Asp | Ala | L;:s | | 590 Asp | Arg | Asn |
| | Phe | Val 610 | | Val | Asp | Asp | | 600 Glu | туг | Gln | Thr | | 605 Asn | Phe | Asn | Arg |
| 75 | Pro 625 | | Trp | Leu | Glγ | Gln 630 | 615 Ile | Ala | Tyr | Ala | Сув 635 | 620 Asn | Arg | C.'.e | Net | Ser 640 |
| | | | | | | | | | | | ~ • | | | | | - 4 U |

| | Val | Gly | e12. | Asn | Pro 645 | Asp | Gly | Lys | Asp | Thr 650 | Pro | Gln | Thr | Thr | Asp 655 | Asp |
|------------|----------|------------|------------|------------|------------|--------|------------|------------|------------|------------------|--------|------------|------------|------------|------------|------|
| | Thr | Ile | Ile | 11e 660 | | Thr | His | Ala | Asp 665 | Glu | Thr | Gly | Ala | Lys 670 | | G1n |
| 5 | Thr | Leu | G1; 675 | | Ala | Ala | Glu | Asn 680 | | Val | Leu | Glu | Glu 685 | Ile | T7.r | Val |
| | Ser | Leu 690 | | Pro | Lys | Ala | Asn 695 | Ser | Ala | Vāl | His | Ala 700 | Leu | Asn | Туг | Leu |
| 10 | 705 | | _ | | _ | 710 | | | | Ser | 715 | | | | | 720 |
| | | | | | 725 | | | | | Glu 730 | | | | | 735 | |
| | | | | 740 | | | | | 745 | Leu | | | | 750 | | |
| 15 | | | 755 | | | | | 760 | | Leu | | | 765 | | | |
| | Gln | Thr 770 | Glu | Leu | Asn | Gln | Leu 775 | Ser | Leu | Arg | Gly. | Glu 780 | Val | Leu | Tyr | Туг |
| 20 | 785 | | | | _ | 790 | | | | Val | 795 | | | | | 800 |
| | Phe | Val | Gln | Het | Ile 805 | His | G1 y | Glu | Ile | Leu 810 | Gln | Lys | Asp | Asn | 11e 815 | Asn |
| | Arg | G1?. | Thr | Val 820 | Pro | Lys | Asp | lle | Phe 825 | Glu | C;;'\$ | Lys | Leu | His 830 | Asn | Leu |
| 25 | Ser | Ser | G15 835 | Ser | lle | Phe | Glu | Glu 840 | Asp | G1; | Gln | Asn | Gly 845 | Asn | llet | Ile |
| | | 850 | | | | | 855 | | | Val | | 860 | | | | |
| 30 | 865 | | | | | 870 | | | | His | 875 | | | | | 880 |
| | | | | | 885 | | | | | Pro 890 | | | | | 895 | |
| | | | | 900 | | | | | 905 | Asn | | | | 910 | | |
| 35 | | | 915 | | | | | 920 | | Туr | | | 925 | | | |
| | | 930 | | | | | 935 | | | Gln | | 940 | | | | |
| 40 | 945 | | | | | 950 | | | | Ala | 955 | | | | | 960 |
| | | | | | 965 | | | | | Ile 970 | | | | | 975 | |
| 45 | | | | 980 | | | | | 985 | | | | | 990 | | |
| 45 | | | 995 | | | _ | | 100 | 0 | Asn | | | 10 | 05 | | |
| | | 101 | 0 | | | | 10 | 15 | | Thi | | 1 | 020 | | | |
| 50 | 102 | 5 | | | | 10 | 30 | | | Asp | 1 | 035 | | | | 1040 |
| | | | | | 104 | 5 | | | | His 10 | 50 | | | | 1 | 055 |
| 55 | | | | 106 | 0 | | | | 10 | | | | | 1 | 070 | |
| 33 | | | 107 | 5 | | | | 10 | 80 | Val | | | 1. | 085 | | |
| | | 109 | 0 | | | | 10 | 95 | | Asn | | 1 | 100 | | | |
| 60 | 110 | 5 | | | _ | 11 | 10 | | - | Gln | 1 | 115 | | | | 1120 |
| | | | | | 112 | 5 | | | | Asn 11 Pro | 30 | | | | 1 | 135 |
| 65 | | | | 114 | 0 | | | | 11 | 45 Ile | | | | 1 | 150 | |
| 03 | _ | _ | 115 | 5 | | | | 11 | 60 | His | | | 1 | 165 | | |
| | - | 117 | 0 | | • | - | 11 | 75 | | | | 1 | 180 | | | |
| 70 | 118 | 5 | | | | 11 | 90 | | | Leu Tyr | 1 | 195 | | | | 1200 |
| | | | | | 120 | 5 | | | | | 10 | | | | 1 | 215 |
| <i>7</i> 5 | | | | 122 | 0 | _ | | _ | 12 | 25 Asn | | | | 1 | 230 | |
| , , | 1. T. C. | TIIT | VIG | 26[| UTa | LI T E | vr ð | nsμ | GIU | U=1) | тrb | Aat | 111 a | • 01 | v a I | |

WO 99/29870 PCT/AU98/01023

```
1240
                                                                1245
       Glu the Lys Glu Lys Ile Leu Arg Ile Thr Lys Gln Glu Val Asn Thr
         1250
                                    1255
       Asp Glu
 5
       1265
       (2) INFORMATION FOR SEQ ID NO:516
             (i) SEQUENCE CHARACTERISTICS:
10
                   (A) LENGTH: 1232 amino acids
                   (B) TYFE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
15
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Perphyromenas gingivalis
20
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...1232
25
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:516
       Het Thr Thr Lys Lys Pro Gln Ala Ile Leu Asp Leu Glu Lys Ala Tyr
1 5 10 15
       Asn Ile Glu Ile Pro Asp Leu Ser Ser Gln Glu Gly Ile Ser Trp Ser 25 30
30
       Val Asn Arg Tyr Phe Lys Gln Asp Ser Ser Gly Ala Val Val Glu Leu
35 40
       Cys Leu Arg Glu Cys Gln Ile Glu Ser Het Thr Trp Leu Ile Asp Phe
50 55 60
35
       Pro Ala Leu Lys Lys Leu Asp Leu Ser Tyr Asn Gln Ile Ser Lys Leu
65 70 75 80
       Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr Lys Leu Arg Leu Arg Ser
85 90
       Asn Gln Ile Arg Lys Leu Glu Gly Leu Asp Ser Leu Thr Ser Leu Thr 100 105 110
40
       Lys Leu Ser Leu Ser Asp Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu
115 120 125
       Arg Leu Thr Ser Leu Ala Glu Leu Tyr Leu Leu Asp Asn Gln Ile Ser
130 135 140
       Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu
145 150 160
45
       Ser Gl; Asn Gln Ile Arg Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser
165 170 175
       Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu Gly
180 185 190
50
       Leu Glu Arg Leu Thr Ser Leu Thr Lys Leu Arg Leu Arg Ser Asn Gln
195 200 205
       Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu 210 215 220
       Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu Gly Leu Glu Arg Leu
225 230 235 240
55
       Thr Ser Leu Ala Thr Leu Glu Leu Ser Gl\gamma Asn Gln Ile Ser Lys Leu 245 250 255
       Glu Gly Leu Glu Arg Leu Ser Ser Leu Thr Lys Leu Arg Leu Arg Ser 260 265 270
60
       Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr 275 280 285
       Lys Leu Ser Leu Ser Asp Ash Gln Ile Ser Lys Leu Glu Gly Leu Glu
290 295 300
65
       Arg Leu Thr Ser Leu Ala Glu Leu Tyr Leu Leu Asp Asn Gln Ile Arg
305 310 315
       Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr Lys Leu Arg Leu 325 330 335
       Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu Asp Ser Leu Thr Ser 340 345 350
70
       Leu Thr Lys Leu Ser Leu Ser Asp Asn Gln Ile Ser Lys Leu Glu Gly 355 360 365
       Leu Glu Arg Leu Thr Ser Leu Ala Glu Leu Tyr Leu Leu Asp Asn Gln
370 375 380
75
       Ile Arg Lys Leu Glu Gly Leu Asp Gly Leu Ala Ser Leu Thr Arg Leu
```

PCT/AU98/01023

| | 205 | | | | | 390 | | | | | 395 | | | | | 400 |
|------------|------------|------------|-------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|
| | 385 Ser | Leu | Arg | Arg | Asn 405 | Gln | Ile | Ser | Lys | Leu 410 | Glu | Gl y | Leu | Asp | Arg 415 | |
| 5 | Lys | Val | Leu | Arg 420 | | Leu | Asp | Val | Ser 425 | | Asn | Asp | Ile | Gln 430 | | Ile |
| J | Asp | Asp | Tle 435 | | Leu | Leu | Ala | Pro 440 | | Leu | Glu | Gln | Thr 445 | | Glu | Lys |
| | Leu | Arg 450 | Ile | His | Asp | Asn | Pro 455 | Phe | Vāl | Ala | Ser | Ser 460 | Gly | Leu | Ile | Leu |
| 10 | 465 | | Tyr | | | 470 | | | | | 475 | | | | | 480 |
| | | | Glu | | 485 | | | | | 490 | | | | | 495 | |
| 15 | | | Иet | 500 | | | | | 505 | | | | | 510 | | |
| | | | Tyr 515 | | | | | 520 | | | | | 525 | | | |
| • | | 530 | His | - | | | 535 | | | | | 540 | | | | |
| 20 | 545 | | Asp | - | _ | 550 | | | | | 555 | | | | | 560 |
| | | | T7.r | | 565 | | _ | | | 570 | | | | | 575 | |
| 25 | | | Asp | 580 | | | | | 585 | | | | | 590 | | |
| | | | 51 y 595 | | | | | 600 | | | | | 605 | | | |
| 30 | | 610 | Pro Gln | | | | 615 | | | | | 620 | | | | |
| 30 | 625 | | Ala | | | 630 | | | | | 635 | | | | | 640 |
| | | | Lys | | 645 | | | | | 650 | | | | | 655 | |
| 35 | | | Arg | 660 | | | | | 665 | | | | | 670 | | |
| | | | 675 Asp | | | | | 680 | | | | | 685 | | | |
| 40 | | 690 | His | | | | 695 | | | | | 700 | | | | |
| | 705 | - | Arg | | | 73.0 | | | | | 715 | | | | | 720 |
| | | | Asn | | 725 | | | | | 730 | | | | | 735 | |
| 45 | | | Lλ.ε | 740 | | | | | 745 | | | | | 750 | | |
| | Gln | Het | 755 Ile | His | Gly | Glu | | 760 Leu | Gln | Lys | Asp | Asn | 765 Ile | Asn | Arg | Gly |
| 50 | Thr | 770 Val | Pro | Lys | Asp | Ile | 775 Phe | Glu | Cys | Lys | | 780 His | Asn | Leu | Ser | |
| | 785 Gly | | Ile | Phe | Glu | 790 Glu | Asp | Gly | Gln | | 795 Gly | Asn | Met | Ile | | 800 Gln |
| | Leu | Leu | Leu | G1 u | 805 Glu | Leu | Ile | Val | Tyr | 810 Glu | Asp | Lγε | Asp | Сув | 815 Tyr | Val |
| 55 | Ile | Pro | Gly | 820 Tyr | | Pro | Leu | His | | | Asp | Glu | | | Γλ.ε | Trp |
| | Leu | | 835 Leu | Gly | Phe | Glu | | | | Phe | Val | | | | Glu | Arg |
| 60 | Phe 865 | | Pro | Phe | Gly | Leu 870 | 855 Ile | | Gln | Ile | Ile 875 | | | Tyr | G17 | Arg 880 |
| | | | Gl; | Ala | Leu 885 | | Arg | Tyr | Trp | Arg 890 | Asp | | Val | lle | Phe 895 | |
| 65 | Ālā | Gl y | Arg | Glu 900 | Het | Asp | Arg | Gln | Thr 905 | Leu | | Gln | Glu | Glu 910 | Glu | $L\gamma s$ |
| | Glu | Gly | Leu 915 | | | Thr | Asn | Ala 920 | Glu | | Туг | Gln | 11e 925 | Trp | | Γλε |
| | Leu | Asp 930 | Phe | Thr | Asp | Leu | Ala 935 | Ile | | Val | Phe | Ile 940 | Lys | | Gln | Arg |
| 70 | Lys 945 | Thr | Ser | Ala | Lys | Asp 950 | Het | | Arg | Lys | Glu 955 | | Thr | Ile | Leu | Ser 960 |
| | Asp | Het | Leu | Asp | 11et 965 | _ | Trp | Asn | Asn | 11e 970 | | Pro | Arg | Glu | Gln 975 | |
| <i>7</i> 5 | Gly | Asp | Lys | Asp 980 | | Glu | Gln | Thr | Arg 985 | | Thr | Ile | Arg | Glu 990 | | Asn |
| | | | | | | | | | | | | | | | | |

```
Arg Lys Lys Arg Fro Ile Gln Asp Leu Tyr Leu Ser Cys Ala Gln Ala 995 1000 1005 1005 Asp Lys Asp Leu Thr Glu Ser His Tyr Ile His Leu Gly Thr Leu Asp 1010 1015 1020
        Asp Glu Ser Lys Thr Thr Ala Arg Ile Ala Ala Tyr Pro Leu Lys Asn
1025 1030 1035
        Gly Val Ile Asp Lys Glu Arg Val Arg Glu Val Ser Thr Arg Pro Tyr
1045 1050 1055
        Lys His Leu Ser Val Asn Lys Asn Leu Ala Thr Ala Lys Gln Ile Phe 1060 1065 1070
10
        Ile Ser Tyr Ser Lys Glu Asp Gln Thr Glu Leu Glu Thr Cys Leu Gln
1075 1080 1085
         Phe Phe Lys Pro Leu Glu Lys Asn Gly Gln Ile Glu Ile Tyr Tyr Asp
1090 1095 1100
        Lys Leu Thr Lys Phe Glu Thr Pro Ile His Pro Glu Ile Arg Lys Arg
1105 1110 1115
15
        Ile Val Glu Ala Asp Cys Ile Ile Ala Leu Ile Ser Gln Arg Tyr Leu
1125 1130 1135
        Ala Thr Asp Tyr Ile Leu Asp His Glu Leu Pro Val Phe Arg Glu Tyr 1140 1145 1150
20
        Asn Lys Thr Ile Val Pro Ile Leu Ile Lys Pro Cys Thr Phe Glu Asp
1155 1160 1165
        Asp Glu Phe Leu Arg Glu Lys Tyr Phe Ala Gln Lys Ala Gln Ile Ile 1170 1175 1180

Asn Leu Gly Lys Glu Gly Lys Thr Ile Lys Ala Tyr Asp Ser Ile Thr 1185 1190 1195
25
        Ala Ser Ala His Arg Asp Glu Asn Trp Val Ala Val Val Arg Glu Phe
1205 1210 1215
Lys Glu Lys Ile Leu Arg Ile Thr Lys Gln Glu Val Asn Thr Asp Glu
30
                       1220
        (2) INFORMATION FOR SEQ ID NO:517
               (i) SEQUENCE CHARACTERISTICS:
35
                     (A) LENGTH: 1175 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
             (ii) NOLECULE TYPE: protein
40
            (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                     (A) ORGANISM: Porphyromonas gingivalis
45
             (ix) FEATURE:
                     (A) NAME/KEY: misc_feature
                     (B) LOCATION 1...1175
50
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:517
       Net Thr Trp Leu Ile Asp Phe Pro Ala Leu Lys Lys Leu Asp Leu Ser \frac{1}{10} \frac{15}{15}
        Tyr Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu
20 25 30
55
        Thr Lys Leu Arg Leu Arg Ser Asn Gln Ile Arg Lys Leu Glu Gly Leu 35 40 45
        Asp Ser Leu Thr Ser Leu Thr Lys Leu Ser Leu Ser Asp Asn Gln Ile 50 55 60
60
        Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Glu Leu Tyr
65 70 75 80
        Leu Leu Asp Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr
85 90 95
        Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu 100 105 110
65
        Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn
115 120 125
        Gln Ile Arg Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr Lys
130 135 140
                                   135
                                                             140
70
        Leu Arg Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg
145 150 155 160
       Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys
165 170 175
        Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser
75
                                              185
```

| | Gly | Asn | Gln 195 | Ile | Ser | Lys | Leu | Glu 200 | Gly | Leu | Glu | Arg | Leu 205 | Ser | Ser | Leu |
|----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Thr | Lys 210 | Leu | Arg | Leu | Arg | Ser 215 | | Gln | Ile | Ser | Lys 220 | | Glu | Gly | Leu |
| 5 | Glu 225 | | Leu | Thr | Ser | Leu 230 | | ГÀг | Leu | Ser | Leu 235 | | Asp | Asn | Gln | 11∈ 240 |
| | Ser | rλæ | Leu | Glu | Gly 245 | Leu | Glu | Arg | Leu | Thr 250 | | Leu | Alā | Glu | Leu 255 | Туг |
| 10 | Leu | Leu | Asp | Asn 260 | Gln | lle | Arg | Lys | Leu 265 | | Gly | Leu | Glu | Arg 270 | Leu | Thr |
| • | Ser | Leu | Thr 275 | | Leu | Arg | Leu | Arg 280 | | Asn | Gln | Ile | Ser 285 | Lys | Leu | Glu |
| | Gly | Leu 290 | Asp | Ser | Leu | Thr | Ser 295 | Leu | Thr | Lys | Leu | Ser 300 | Leu | Ser | Asp | Asn |
| 15 | 305 | | | | | 310 | | | | | 315 | | | | Alā | 320 |
| | | | | | 325 | | | | | 330 | | | | | Asp 335 | |
| 20 | | | | 340 | | | | | 345 | | | | | 350 | Ser | |
| | | | 355 | | | | | 360 | | | | | 365 | | Val | |
| | | 370 | | | | | 375 | | | | | 380 | | | Pro | |
| 25 | 385 | | | | | 390 | | | | | 395 | | | | Phe | 400 |
| | | | | | 405 | | | | | 410 | | | | | Pro 415 | |
| 30 | | | | 420 | | | | | 425 | | | | | 430 | Thr | |
| | | | 435 | | | | | 440 | | | | | 445 | | His | |
| 25 | | 450 | | | | | 455 | | | | | 460 | | | Thr | |
| 35 | 465 | | | | | 470 | | | | | 475 | | | | Pro | 480 |
| | | | | | 485 | | | | | 490 | | | | | Ile 495 | |
| 40 | | | | 500 | | | | | 505 | | | | | 510 | Asp Gln | |
| | | | 515 | | | | | 520 | | | | | 525 | | Ala | |
| 45 | | 530 | | | | | 535 | | | | | 540 | | | Thr | |
| 10 | 545 | | | | | 550 | | | | | 555 | | | | Glu | 560 |
| | | | | | 565 | | | | | 570 | | | | | 575 Val | |
| 50 | | | | 580 | | | | | 585 | | | | | 590 | Val | |
| | | | 595 | | | | | 600 | | | | | 605 | | Ser | |
| 55 | | 610 | | _ | | | 615 | _ | | | | 620 | | | Glu | |
| | 625 | | | | | 630 | | | | | 635 | | | | Leu | 640 |
| | | | | | 645 | | | | | 650 | | | | | 655 Leu | |
| 60 | Thr | Ile | Glu | 660 Tyr | | Gln | Thr | Glu | 665 Leu | | Gln | Leu | | 670 Leu | Arg | G1 y |
| | Glu | Val | 675 Leu | Tyr | T;·r | Arg | | | | L;∵s | Leu | | | T;r | Val | Trp |
| 65 | | | | Ala | Ala | | | | Het | Ile | | | | Ile | Leu | |
| | 705 Lys | | Asn | Ile | | 710 Arg | | Thr | Val | | | | Il∈ | Phe | Glu | 720 Cys |
| 70 | Lys | Leu | His | | | Ser | Ser | Gly | | | | Glu | Glu | Asp | 735 Gly | |
| 70 | Asn | Gly | | | | Leu | Gln | | | | Glu | Glu | | | Val | Tyr |
| | Glu | | | | Сув | Tyr | | | | G1 ? | туг | | | | His | Ser |
| 75 | Asp | 770 Asp | | Ala | Туг | Lys | 775 Trp | | Thr | Leu | Gly | 780 Phe | | Arg | Pro | Asn |

```
790
                                                       795
       Phe Val Leu Lys Phe Glu Arg Phe Ile Pro Phe Gly Leu Ile Asn Gln
805 810 815
       Ile Ile Ala Tyr Tyr Gly Arg Glu Glu Gly Ala Leu Lys Arg Tyr Trp
820 825 830
       Arg Asp Gln Val Ile Fhe Thr Ala Gly Arg Glu Het Asp Arg Gln Thr 835 840 845
       Leu Glu Glu Glu Glu Lys Glu Gly Leu Pro Lys Thr Asn Ala Glu
850 855 860
10
       Asp Tyr Gln Ile Trp Ile Lys Leu Asp Phe Thr Asp Leu Ala Ile Ser 865 870 875
       Val Phe Ile Lys Glu Gln Arg Lys Thr Ser Ala Lys Asp Met Gln Arg
885 890 895
       Lys Glu Ala Thr Ile Leu Ser Asp Het Leu Asp Het Tyr Trp Asn Asn 900 905 910
15
       lle Pro Pro Arg Glu Gln Ile Gly Asp Lys Asp Thr Glu Gln Thr Arg
915 920 925
       Ser Thr Ile Arg Glu Thr Asn Arg Lys Lys Arg Pro Ile Gln Asp Leu
930 935 940
20
       Tyr Leu Ser Cys Ala Gln Ala Asp Lys Asp Leu Thr Glu Ser His Tyr
945 950 955 960
                                                   955
       lle His Leu Gly Thr Leu Asp Asp Glu Ser Lys Thr Thr Ala Arg Ile
965 970 975
       Ala Ala Tyr Pro Leu Lys Ash Gly Val Ile Asp Lys Glu Arg Val Arg
980 985 990
25
       Glu Val Ser Thr Arg Pro Tyr Lys His Leu Ser Val Asn Lys Asn Leu
995 1000 1005
       Ala Thr Ala Lys Gln Ile Phe Ile Ser Tyr Ser Lys Glu Asp Gln Thr
1010 1015 1020
30
       Glu Leu Glu Thr Cys Leu Gln Phe Phe Lys Pro Leu Glu Lys Asn Gly
1025 1030 1035
                                                        1035
       Gln Ile Glu Ile Tyr Tyr Asp Lys Leu Thr Lys Phe Glu Thr Pro Ile
1045 1050 1055
       His Pro Glu Ile Arg Lys Arg Ile Val Glu Ala Asp Cys Ile Ile Ala 1060 1065 1070

Leu Ile Ser Gln Arg Tyr Leu Ala Thr Asp Tyr Ile Leu Asp His Glu 1075 1080 1085
35
       Leu Pro Val Phe Arg Glu Tyr Asn Lys Thr Ile Val Pro Ile Leu Ile
1090 1095
40
       Lys Pro Cys Thr Phe Glu Asp Asp Glu Phe Leu Arg Glu Lys Tyr Phe
1105 1115
       Ala Gln Lys Ala Gln Ile Ile Asn Leu Gly Lys Glu Gly Lys Thr Ile
1125 1130 1135
       Lys Ala Tyr Asp Ser Ile Thr Ala Ser Ala His Arg Asp Glu Asn Trp
1140 1145 1150
       Lys Ald Tyl hep 602 42.

1140 1145 1160 1160

Val Ala Val Val Arg Glu Phe Lys Glu Lys Ile Leu Arg Ile Thr Lys
45
       Gln Glu Val Asn Thr Asp Glu
           1170
50
       (2) INFORMATION FOR SEQ ID NO:518
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 229 amino acids
55
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) HOLECULE TYPE: protein
60
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISH: Porphyromonas gingivalis
65
            (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                   (B) LOCATION 1...229
            (Xi) SEQUENCE DESCRIPTION: SEQ ID NO:518
70
       Het Het Lys Lys Ala Phe Val Phe Val Leu Leu Val Cys Leu Phe Ser
                                                  10
       Ser Phe Ser Ser Ser Ala Gln Thr Thr Thr Asn Ser Ser Arg Ser Tyr
                                             25
                                                                     30
75
       Phe Thr Gly Arg Ile Glu Lys Val Ser Leu Asn Leu Gly Val Pro Pro
```

```
40
                 35
       Val Ser Thr Glu Val Trp Gly Net Thr His Asp Ala Ash Gly Leu Pro
50 55 60
       Phe Glu Ile Pro Ile Ser Phe Ser Arg Phe Asn Ser Gln Gly Asp Ile
65 70 75 80
 5
       Ala Thr Thr Tyr Tyr Ile Ala Asn Ser Glu Ala Thr Leu Asn Glu Trp
85 90 95
       Cys Asp Tyr Ala His Pro Gly Gly Ile Val Arg Val Glu Gly Arg Phe
100 105 110
10
       Trp Lys Het Thr Tyr Asn Ile Pro Thr Tyr Asn Ala Val Cys Thr Arg 115 120 125
       Ile Thr Phe Glu Asn Gln Glu Ile Glu Gly Thr Ile Val Leu Ile Pro
130 135 140
       Lys Pro Lys Val Ser Leu Pro His Val Ser Glu Ser Val Pro Cys Ile
145 150 150 160
15
       Arg Thr Glu Ala Gly Arg Glu Phe Ile Leu Cys Glu Glu Asp Asp Thr
165 170 175
       Phe Val Ser His Asp Gly Asn Glu Val Thr Ile Gly Gly Lys Pro Phe
180 185 190
20
       Leu Leu Asn Thr Asn Val Lys Ile Val Gly Asp Val Ser Gln Lys Tyr
195 200 205
       Ala Val Gly Val Gly Glu Ile Arg Phe Leu Gln Ile Cys Ala Gln Thr
       Val Ser Gln Gln Lys
25
       (2) INFORMATION FOR SEQ ID NO:519
             (i) SEQUENCE CHARACTERISTICS:
30
                   (A) LENGTH: 228 amino acids
                   (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
35
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
40
            (ix) FEATURE:
                   (A) NAHE/KEY: misc_feature (B) LOCATION 1...228
45
            (Mi) SEQUENCE DESCRIPTION: SEQ ID NO:519
       Het Lys Lys Ala Phe Val Phe Val Leu Leu Val Cys Leu Phe Ser Ser
       Phe Ser Ser Ser Ala Gln Thr Thr Thr Asn Ser Ser Arg Ser Tyr Phe
20 25 30
50
       Thr Gly Arg Ile Glu Lys Val Ser Leu Asn Leu Gly Val Pro Pro Val 35 40 45
       Ser Thr Glu Val Trp Gly Het Thr His Asp Ala Asn Gly Leu Pro Phe 50 60
       Glu Ile Pro Ile Ser Phe Ser Arg Phe Asn Ser Gln Gly Asp Ile Ala
65 70 75 80
55
       Thr Thr Tyr Tyr Ile Ala Asn Ser Glu Ala Thr Leu Asn Glu Trp Cys
85 90 95
       Asp Tyr Ala His Pro Gly Gly Ile Val Arg Val Glu Gly Arg Phe Trp
100 105 110
60
       Lys Het Thr Tyr Asn Ile Pro Thr Tyr Asn Ala Val Cys Thr Arg Ile
115 120 125
       Thr Phe Glu Asn Gln Glu Ile Glu Gly Thr Ile Val Leu Ile Pro Lys
130 135 140
       Pro Lys Val Ser Leu Pro His Val Ser Glu Ser Val Pro Cys Ile Arg
145 150 155 160
65
       Thr Glu Ala Gly Arg Glu Phe Ile Leu Cys Glu Glu Asp Asp Thr Phe
165 170 175
       Val Ser His Asp Gly Asn Glu Val Thr Ile Gly Gly Lys Pro Phe Leu
180 185 190

Leu Asn Thr Asn Val Lys Ile Val Gly Asp Val Ser Gln Lys Tyr Ala
195 200 205
70
       Val Gly Val Gly Glu Ile Arg Phe Leu Gln Ile Cys Ala Gln Thr Val
210 215 220
          210
                                  215
75
       Ser Gln Gln Lys
```

WO 99/29870 PCT/AU98/01023

474/490

225

```
(2) INFORMATION FOR SEQ ID NO:520
 5
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 540 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
10
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
15
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
                   (A) MANIE/KEY: misc_feature
                   (B) LOCATION 1...5\overline{4}0
20
            (::i) SEQUENCE DESCRIPTION: SEQ ID NO:520
       Het Lys Thr Lys Val Leu Arg Lys Phe Val Val Ala Ala Phe Ala Val
25
       Ala Thr Leu Cys Pro Leu Ala Gln Ala Gln Thr Het Gly Gly Asp Asp
20 25 30
       Val Lys Val Val Gln Tyr Asn Gln Glu Lys Leu Val Gln Thr Arg Met 35 40
       Ser Val Ala Asp Asn Gly Trp Ile Tyr Val Het Thr His Ser Gly Tyr
50 60
30
       Asp Thr Gl\gamma Asn Ser Asn Val Lys Ile Phe Arg Ser Lys Asp Gln Gly G5 70 75
       Ala Thr Tyr Gln Lys Leu Arg Asp Trp Asp Pro Ser Asp Asp Tyr Gln 85 \\ 90 \\ 95
       Phe Gln Asp Phe Asp Ile Val Val Thr Gly Lys Asn Glu Ser Asp Ile
100 105 110
35
       Lys lle Trp Ser Val Glu Leu Het Asn Lys Pro Gly Gly Tyr Lys Ser 115 120 125
       Arg Val Ala Val Phe Ser Arg Asp Ala Asn Ala Gln Asn Ala Lys Leu
130 135 140
40
       Val Tyr Lys Glu Asp Phe Ser Asn Val Gln Leu Tyr Asp Val Asp Ile
145 150 155 160
       Ala Ser Asn Tyr Arg Ser Pro Ser Ser Leu Asn Asn Gly Gly Asn Pro
165 170 175
       Phe Ala Leu Ala Phe Ala Tyr Thr Gly Phe Asn Asn Thr His Lys Ile
180 185 190
45
       Ser Phe Val Asp Tyr Val Phe Ser Leu Ash Gly Gly Gln Ash Phe Ash
195 200 205
       Lys Asn Leu Leu Phe Ser Gln Asp Gly Glu Lys Lys Ile Asp Lys Val
210 215 220
50
       Asp Leu Ser Leu Gly Ser Thr Ser Glu Ser Het Gly His Asn Ala Trp
225 230 235 240
       Pro Leu Met Gly Val Val Phe Glu Met Asn Lys Gln Gly Gly Lys Ser
245 250 255
55
       Asp Ile Gly Phe Leu Ser Asn Phe Val Asp Asn Asp Pro Glu Phe Gln 260 265 270
       Trp Ser Gly Pro Ile Lys Val Ser Glu Ser Asp Met Ser Phe Ser Pro
275 280 285
       Lys Ile Gln Het Leu Leu Asp Glu Asp Asn Asn Thr Ile Asn Gly Glu 290 295 300

Ser Cys His Asn Phe Met Ile The The Co
60
       Ser Cys His Asn Phe Met Ile Thr Tyr Ser Asp Tyr Asp Ser Glu Tyr
305 310 315 320
       Ser Asp Trp Asp Ile Arg Tyr Val Tyr Pro Lys Lys Ser Phe Lys Tyr 325 330 335
65
       Glu Lys Gly Lys Thr Pro Thr Het Asp Asp Leu Val Glu Ala Phe Leu
340 350
       Thr Ala Ser Tyr Gln Ser Glu Thr Asn Ser Gly Leu Gly Tyr Asp Lys
355 360 365
       Asn Ala Asn His Tyr Leu Ile Thr Tyr Ala Lys Lys Glu Glu Asn Gly
370 375 380
70
       Thr Asn Thr Leu Lys Tyr Arg Trp Ala Asn Tyr Asp Lys Ile His Asn 385 390 395 400
       Lys Asp Leu Trp Ser Asp Thr Phe Thr Tyr Thr Ser Ser Ala Asn Ala 405 410 415
75
       Leu Tyr Thr Pro Gln Val Asp Ile Asn Pro Thr Lys Gly Leu Val Cys
```

```
430
       Trp Ser Trp Val Glu Tyr Leu Pro Gly Lys Arg lle Val Trp Ser Asp
435 440 445
       Thr Gln Trp Thr His Ala Asn Gly Val Glu Asp Ile Val Het Gln Glu
450 450
 5
       Gly Ser Net Lys Leu Tyr Pro Asn Pro Ala Gln Glu Tyr Ala Val Ile
465 470 475 480
       Ser Leu Pro Thr Ala Ala Asn Cys Lys Ala Val Val Tyr Asp Het Gln
485 490 495
       Gly Arg Val Val Ala Glu Ala Ser Phe Ser Gly Asn Glu Tyr Arg Leu
500 510
10
       Asn Val Gln His Leu Ala Lys Gly Thr Tyr Ile Leu Lys Val Val Ser 515 520 525
Asp Thr Glu Arg Phe Val Glu Lys Leu Ile Val Glu 530 530 530
15
        (2) INFORMATION FOR SEQ ID NO:521
              (i) SEQUENCE CHARACTERISTICS:
20
                    (A) LEHGTH: 771 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
25
           (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) ORGANISH: Porphyromonas gingivalis
30
            (ix) FEATURE:
                    (A) NAME/KEY: misc_feature
                    (B) LOCATION 1...771
35
             (mi) SEQUENCE DESCRIPTION: SEQ ID NO:521
       Het Leu Thr Ile Arg Asn Phe Leu Leu Phe Cys Cys Leu Ser Leu Ile 1 \phantom{000}5\phantom{000} 15
       Ala Phe Ala Ala Asp Ala Gln Ser Ser Val Ser Ser Gly Arg Arg Leu
20 25 30
40
        Thr Glu Tyr Val Asn Pro Phe Ile Gly Thr Ala Asn Tyr Gly Thr Thr 35 40
        Asn Pro Gly Ala Val Leu Pro Asn Gly Leu Met Ser Val Thr Pro Phe
50 60
        Asn Val Ser Gly Ser Thr Glu Asn Arg Phe Asp Lys Asp Ser Arg Trp 65 70 75
45
        Trp Ser Ala Pro Tyr Ser Ala Asp Asn Ser Tyr Cys Ile Gly Phe Ser
85 99
        His Val Asn Leu Ser Gly Val Gly Cys Pro Glu Leu Ser Gly Ile Leu
100 105 110
50
        Leu Het Ala Thr Ser Gly Thr Phe Asp Pro Asp Tyr Cys Cys Tyr Gly
115 120 125
        Ser Ser Leu Ser Arg Glu Tyr Ala Arg Pro Gly Glu Tyr Lys Ala Val
130
135
140
Leu Asp Lys Tyr Gly Ile Asp Ala Ala Val Thr Val Thr Glu Arg Thr
145
150
155
160
55
        Ala Leu Thr Glu Phe Ala Phe Pro Glu Gly Glu Gly His Ile Leu Leu 165 170 175
        Asn Leu Gly Gln Ala Leu Ser Asn Glu Ser Gly Ala Ser Val Arg Phe
180 185 190
60
        Leu Asn Asp Ser Thr Val Val Gly Ser Arg Leu Het Gly Thr Fhe Cys
195 200 205
        Tyr Asn Pro Gln Ala Val Phe Arg Gln Tyr Pne Val Leu Gln Val Ser
210 215 220
65
        Arg Arg Pro Ile Ser Ala Gly Tyr Trp Lys Lys Gln Pro Pro Met Thr 225 230 235 240
        Val Glu Ala Gln Trp Asp Ser Thr Ala Gly Lys Tyr Lys Gln Tyr Asp
245 250 255
        Gly Tyr Lys Arg Glu Net Ser Gly Asp Asp Ile Gly Val Arg Phe Ser 260 270
70
        Phe Asn Cys Asp Gln Gly Glu Lys Ile Tyr Val Arg Ser Ala Val Ser 275 280 285
        Phe Val Ser Glu Ala Asn Ala Leu Tyr Asn Leu Glu Ala Glu Gln Glu
290 295 300
75
        Glu Val Phe Lys Ser Val Gly Gly Asn Pro Ala Lys Ala Phe Ser Ala
```

```
310
        Ile Arg Ser Arg Ala Ile Glu Arg Trp Glu Glu Ala Leu Gly Thr Val
325 330 335
        Glu Val Glu Gly Gly Thr Pro Asp Glu Lys Thr Ile Phe Tyr Thr Ala
340 345 350
 5
        Leu Tyr His Leu Leu Ile His Pro Asn Ile Leu Gln Asp Ala Asn Gly
355 360 365
        Glu Tyr Pro Net Net Gly Ser Gly Lys Thr Gly Asn Thr Ala His Asp
370 375
10
        Leu Leu Cys Leu Leu Tyr Pro Glu Lys Gln Leu Asp Met Val Arg Thr 405 \hspace{1cm} 410 \hspace{1cm} 415
        Leu Ile Asp Met Tyr Arg Glu Ser Gly Trp Leu Pro Arg Trp Glu Leu 420 425 430
15
        Tyr Gly Gln Glu Thr Leu Thr Met Glu Gly Asp Pro Ser Leu Ile Val 435 440 445
        Ile Asn Asp Thr Trp Gln Arg Gly Leu Arg Ala Phe Asp Thr Ala Thr
450 455 460
20
        Ala Tyr Glu Ala Met Lys Lys Asn Ala Ser Ser Ala Gly Ala Thr His
465 470 480
        Pro Ile Arg Pro Asp Asp Asp Tyr Leu Thr Leu Gly Phe Val Pro 485 490 495
        Leu Arg Glu Gln Tyr Asp Asn Ser Val Ser His Ala Leu Glu Tyr Tyr
500 505 510
25
        Leu Ala Asp Trp Asn Leu Ser Arg Phe Ala His Ala Leu Gly His Lys 525 520 525
        Glu Asp Ala Ala Leu Phe Gly Lys Arg Ser Leu Gly Tyr Arg His Tyr
530 535
30
        Tyr Asn Lys Glu Tyr Gly Het Leu Cys Pro Leu Leu Pro Asp Gly Ser
545 550 560

        Phe Leu Thr
        Pro
        Phe Asp Pro Lys
        Gln
        Gly
        Glu
        Asn Phe Glu
        Pro Asn 575

        Pro
        Gly
        Phe His
        Glu
        Gly
        Ser Ala Tyr
        Asn Tyr
        Ala Phe Phe Val
        Pro 580
        590

35
        His Asp Ile Gln Gly Leu Ala Arg Leu Met Gly Gly Ala Lys Val Phe
595 600 605
        Ser Glu Arg Leu Gln Lys Val Phe Asp Glu Gly Tyr Tyr Asp Pro Thr
610 615 620
40
        Asn Glu Pro Asp Ile Ala Tyr Pro Tyr Leu Phe Ser Tyr Phe Pro Lys
625 630 635 640
        Glu Ala Trp Arg Thr Gln Lys Leu Thr Arg Glu Leu Ile Asp Lys His
645 650 655
        Phe Cys Asn Ala Pro Asn Gly Leu Pro Gly Asn Asp Asp Ala Gly Thr
660 665 670
45
       Het Ser Ala Trp Leu Val Tyr Ser Met Leu Gly Phe Tyr Pro Asp Cys
675 680 685
        Pro Gly Ser Pro Thr Tyr Thr Leu Thr Ser Pro Val Phe Pro Arg Val
690 695 700
50
        Arg Ile Arg Leu Asn Pro Gln Tyr Tyr Pro Gln Gly Glu Leu Ile Ile
705 710 715
       Thr Thr Asn Thr Glu Asn Gln Pro Thr Asp Ser Ile Tyr Ile His Thr 725 730 735
        Val Ser Leu Gly Asn Lys Thr Leu Pro His Gly Thr Arg His Ile Ser
740 745 750
55
        His Ala Asp Leu Val Arg Cys Gly His Leu Arg Tyr Glu Leu Ser Asn
755 760 765
        Arg Pro Arg
            770
60
        (2) INFORMATION FOR SEQ ID NO:522
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 776 amino acids
65
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (ii) NOLECULE TYPE: protein
70
           (iii) HYPOTHETICAL: YES
             (vi) ORIGINAL SOURCE:
                    (A) \mathsf{ORGAMISM}: \mathsf{Porphyromonas} \mathsf{gingivalis}
75
             (ix) FEATURE:
```

(A) NAME/KEY: misc_feature (B) LOCATION 1...776

| 5 | | (xi) | SEÇ | ONEHO | CE DE | SCRI | PTIC | DH: 5 | SEQ I | D 110 | 523 | 2 | | | | |
|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------------------|
| J | llet 1 | Cha | Γ'.ε | Ile | Arg | Phe | Ser | Leu | Leu | Gln 10 | Ala | Leu | Val | Vāl | Cys 15 | Leu |
| | Leu | Phe | Thr | Ser 20 | Phe | Ser | Leu | Gln | | | Glu | Glu | Gly | Ile 30 | | Asn |
| 10 | Thr | Leu | Leu 35 | | Ile | His | Lys | Thr 40 | 25 Glu | Lys | Ala | Val | Glu 45 | | Pro | Γ?.ε |
| | Lys | Val | | Ala | Val | Ala | Asn 55 | | Val | Leu | T;·r | Ser 60 | | Gly | Lγε | Glu |
| 15 | Ala 65 | Pro | His | Glu | Ala | Lys 70 | | Phe | Asp | Arg | Ile 75 | Ser | G1? | Leu | Ser | Asp 80 |
| | Thr | Ser | Val | Ser | Ser 85 | Ile | Alā | туг | Ser | Glu 90 | Gln | Leu | Lys | Ser | Leu 95 | Val |
| | Ile | Tyr | T';:r | Ala 100 | Ser | Gly | Asn | Ile | Αερ 105 | Il∈ | Leu | Asp | Glu | Ala 110 | Gl y | Arg |
| 20 | Vāl | Thr | Asn 115 | Val | Pro | Ala | Leu | Lys 120 | Asp | Asn | Ile | Asp | Leu 125 | Ile | Asp | Γ ?. ε |
| | Thr | Leu 130 | Aεn | Arg | Leu | Leu | Ile 135 | Val | Gl y | Asn | Arg | Ala 140 | T'.r | Leu | Ala | Gl; |
| 25 | Gly 145 | Phe | Gly | Leu | Ser | Val 150 | Leu | Asp | Vāl | Ala | Glu 155 | Alā | Arg | Ile | Pro | Ala 160 |
| | Thr | Τγr | Ala | Γ.'.ε | Gl; 165 | Thr | Γ?.ε | Val | Thr | Asp 170 | Val | Ala | Lys | Leu | Asp 175 | Asn |
| | Asp | Arg | Leu | Leu 180 | Het | Leu | Γ7.ε | Glu | Gly 185 | Gln | Leu | Phe | Ile | Gly 190 | Lys | Glu |
| 30 | Thr | Asp | Asn 195 | Leu | Gln | Asp | Pro | Ala 200 | Ala | Trp | Thr | Ala | Leu 205 | Ser | Leu | Asn |
| | Leu | Pro 210 | Het | Glγ | Ser | Vāl | Thr 215 | G1y | Leu | G1;· | Ile | Val 220 | Glγ | Glu | Asp | Ile |
| 35 | Cy€ 225 | Phe | Leu | Leu | Ala | Asp 230 | Gly | Arg | Val | Tyr | Val 235 | Alā | Alā | Asn | Gln | Ser 240 |
| | Phe | Glu | Pro | Glu | Leu 245 | Leu | Leu | Ser | Ser | Ser 250 | Ala | Asp | Ser | Arg | Leu 255 | Туг |
| | Vāl | Thr | Asp | Arg 260 | Gly | Leu | Phe | Ile | Cys 265 | Ala | Glu | Asn | Arg | 11e 270 | Туr | Phe |
| 40 | Ile | Glu | Lys 275 | GJ?. | Arg | Γλε | Thr | Thr 280 | Gln | Phe | Pre | lle | Ala 285 | Asp | Val | Leu |
| | G1 ?. | Val 290 | Glγ | Ala | Met | Asn | Glu 295 | Ser | Asn | Thr | Ala | Tyr 300 | Ile | Ala | Leu | Gly |
| 4 5 | 305 | Glu | | | | 310 | | | | | 315 | | | | | 320 |
| | Ala | Net | Pro | Val | Ala 325 | Phe | Asp | Gly | Pro | 330 330 | Asp | Asn | Asp | Phe | Tyr 335 | Glu |
| | Het | Arg | Phe | Ser 340 | His | Gly | Arg | Leu | Tyr 345 | Ala | Ala | Ser | Gl 7. | Leu 350 | Trp | Gly |
| 50 | | Asn | 355 | | | | | 360 | | | | | 365 | | | |
| | | Trp 370 | | | | | 375 | | | | | 380 | | | | |
| 55 | 385 | Phe | | | | 390 | | | | | 395 | | | | | 400 |
| | | Asp | | | 405 | | _ | | • | 410 | | _ | | | 415 | |
| 00 | | Asp | | 420 | | | | | 425 | | | | | 430 | | |
| 60 | Ala | Glu | C;:s 435 | Asn | Pro | Gly | Asp | Ala 440 | Arg | Val | Γ?.ε | Ala | 11e 445 | Alā | Phe | Asp |
| | Asn | Lys 450 | Gly | Asn | Leu | Trp | G1; 455 | Thr | Leu | Gl y | Ala | Val 460 | Gly | Lys | Asn | Ile |
| 65 | Phe 465 | Het | Туг | Asp | Pro | Gln 470 | Ser | Ser | Thr | Trp | His 475 | Ser | Ьре | Ser | Tγr | Pro 480 |
| | | Val | | | 485 | | | | | 490 | | | | | 495 | |
| 70 | | Asp | | 500 | | | | | 505 | | | | | 510 | | |
| 70 | | Gly | 515 | | | | | 520 | _ | _ | | | 525 | | | |
| | _ | Asp 530 | | | | _ | 535 | | | | | 540 | _ | | - | |
| 75 | Ala 545 | Ile | Gl? | His | Lγε | Thr 550 | Ile | ТУ·г | Ala | Met | Ala 555 | Val | Asp | His | Asn | G1; 560 |

```
Ser Val Trp Met Gly Ser Asp Ile Gly Ile Phe Gly Val Tyr Asn Ala
565 570 575
       Ala GJ; Val Leu Ser Ser Thr Ser Thr Pro Ile Ala Val Arg Pro Val 580 585 590
 5
       Gly Gly Glu Glu Pro Asn Leu Tyr Tyr Val Leu Asp Lys Val Thr Val
595 600 605
        Thr Asp Ile Val Val Asp Lys Leu Asn His Lys Trp Val Ala Thr Gln
610 615 620
       Gly Thr Gly Leu Tyr Leu Leu Ser Glu Asp Cys Ser Lys Ile Leu Ala
625 630 635 640
10
       Gln Phe Thr Val Glu Asn Ser Pro Leu Leu Ser Asn Asn Ile Leu Ser
645 650 655
       Leu Ala Leu Asn Asp Asp Asn Gly Leu Leu Tyr Ile Gly Thr Ala Asp 660 665 670
       Gly Leu Met Thr Phe Gln Thr Gly Thr Gly Ser Gly Ser Ala Ser Glu
675 680 685
15
       Leu Asp Gly Val Tyr Val Tyr Pro Asn Pro Leu Arg Pro Glu Tyr Pro
690 695 700
       Asp Gly Val Thr Ile Ala Gly Leu Gln Ala Gly Cys Ser Val Lys Ile
705 710 715
20
       Thr Asp Thr Thr Gly Arg Leu Leu Tyr Gln Thr Glu Ser Val Thr Thr 725 730 735
       Glu Val Lys Trp Asn Ala Arg Gly Ala Asp Gly Asn Arg Val Ala Ser
740 745
25
       Gly Val Tyr Ala Val Ala Val Tyr Asp Pro Val Ser Lys Ser Lys
765 760
       Leu Ile Arg Phe Ala Val Ile Arg
30
       (2) INFORMATION FOR SEQ ID NO:523
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1158 amino acids
                   (B) TYPE: amino acid
35
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
40
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
45
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...1158
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:523
       Het Lys Arg Ile Leu Pro Ile Val Ala Phe Leu Ser Leu Phe Leu Ala 1 \hspace{1.5cm} 5 \hspace{1.5cm} 10 \hspace{1.5cm} 15
50
       Leu Ala Leu Pro Ala Lys Ala Gln Arg Ala Het Gly Lys Thr Ala Asp
20 25 30
       Arg Ser Leu Met Ala Ser Gly His Trp Val Lys Ile Arg Val Asp Ala
35 40 45
55
       Ser Gly Val Tyr Arg Leu Thr Asp Glu Gln Leu Arg Ala Asn Gly Phe
50 55
       Ser Asp Pro Ser Lys Val Gly Val Phe Gly Tyr Gly Gly Gly Val Leu 65 70 75
60
       Pro Glu Asp Leu Ser Arg Ile Thr Thr Asp Asp Leu Pro Pro Val Pro
85 90
       Val Leu Arg Gln Gly Asn Ala Leu Tyr Phe Tyr Ala Val Gly Pro Val
       Thr Trp Fhe Tyr Asn Pro Ala Lys Thr Thr Het Glu His Thr Val Asn 115 120 125
65
       Thr Tyr Ser Thr His Gly Tyr Tyr Phe Leu Ser Asp Ala Ala Gly Ala
130 135 140
       Pro Leu Gln Het Ser Gln Tyr Thr Gly Gly Gly Ala Ser Ala Glu Ala
145 150 160
70
       Leu Ile Asp Tyr Tyr Asp Glu Leu Met Leu His Glu Gln Glu Leu Tyr
165 170 175
       Ser Pro Lys Glu Ser Gly Arg Asp Leu Tyr Gly Glu Ser Phe Ser Ala
180 180
       Val Asn Thr Arg Thr Val Lys Phe Pro Leu Arg Gly Asn Thr Arg Ser
75
                                       200
```

| | | 210 | | | | | 215 | | Ser | | | 220 | | | | |
|------------|------------|------------|------------|-------------|------------|------------|---------------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|
| | Ala 225 | Gly | Gly | Gly | Arg | Glu 230 | Иet | Ser | Leu | Ser | Ala 235 | Asn | Gly | Ile | Leu | 11e 240 |
| 5 | Phe | Ser | Asp | Pro | Phe 245 | Ser | Met | Thr | Ser | Asn 250 | Glu | Val | Ser | Asn | Ser 255 | Туг |
| | Leu | Ala | Gly | 15;€ 260 | | Arg | Arg | Leu | Tyr 265 | | Ser | Thr | Pro | Met 270 | Asn | Ser |
| 10 | Leu | Val | Asn 275 | Glu | Leu | Arg | Leu | Asp 280 | Ala | Asn | Tyr | Ser | Het 285 | Thr | Glγ | Asp |
| | Ala | Val 290 | Asn | Leu | Asp | Phe | Ile 295 | | Vāl | Ala | Thr | Gln 300 | Asn | Asp | Leu | Arg |
| | Tyr 305 | Asp | Gly | Ala | Pro | Met 310 | | Ile | Arg | Arg | Phe 315 | Ser | Asn | Leu | Pro | Val 320 |
| 15 | Leu | G1 7. | Gly | Glu | Ser 325 | C7.2 | Arg | Phe | Val | Ile 330 | | Glu | Val | Pro | Glu 335 | Ser |
| | Leu | Val | Vāl | Leu 340 | | Ala | Asn | Ser | Ser 345 | | Thr | Ala | Ser | Leu 350 | Val | Pro |
| 20 | Val | Lys | The 355 | Val | Gly | Asp | r?.e | Thr 360 | Ile | Glu | Phe | Val | Ala 365 | Pro | Pro | rs |
| | Glγ | Gln 370 | Asp | Arg | Arg | Thr | 11e 375 | Asn | Thr | Phe | Tyr | Ala 380 | Val | Asp | Leu | Ser |
| | Gln 385 | Ala | Ser | Ala | Pro | Glu 390 | Ile | Leu | Glγ | Ala | Val 395 | Pro | Asn | Gln | Asn | Leu 400 |
| 25 | His | Glγ | Glu | Glu | 11e 405 | Pro | Asp | Leu | Ile | 11e 410 | Val | Ser | Thr | Gln | Ala 415 | Leu |
| | Leu | Leu | Glu | Ala 420 | Asp | Arg | Leu | Ala | Thr 425 | Tyr | Arg | Arg | Glu | Lys 430 | Asn | Gl? |
| 30 | Leu | Γλε | Val 435 | Leu | Val | Val | Leu | Gln 440 | Glu | Gln | Val | Phe | Asn 445 | Glu | Phe | Ser |
| | Gl y | G17 450 | Thr | Pro | Asp | Ala | Thr 455 | Ala | Tyr | Arg | Leu | Phe 460 | Ala | Lγε | llet | Phe |
| | Tyr 465 | Asp | Arg | Trp | Lys | Ala 470 | Asn | Ala | Pro | Val | Gly 475 | Glu | Thr | Phe | Pro | Met 480 |
| 35 | Gln | Het | Leu | Leu | Phe 485 | G1 y | Asp | Gly | Ala | His 490 | Asp | Asn | Arg | rla | Val 495 | Ser |
| | Val | Ala | Trp | Gln 500 | r?.s | Pro | Tyr | Leu | Gln 505 | Gln | Thr | Glu | Phe | Leu 510 | Leu | Thr |
| 40 | Phe | Gln | Ala 515 | Vā1 | Asn | Ser | Thr | Asn 520 | Val | Asn | Ser | Туг | Val 525 | Thr | Asp | Asp |
| | туг | Phe 530 | _ | Leu | Leu | Asp | А <i>в</i> р 535 | | Pro | Ala | Ser | Val 540 | Asn | Ile | Gly | Trp |
| | Arg 545 | Asn | Tyr | Asn | liet | Ala 550 | Val | Gly | Arg | Phe | Pro 555 | Val | Arg | Thr | Pro | Ala 560 |
| 45 | Glu | Ala | Arg | Ile | Ala 565 | Val | Asp | Lys | Thr | Ile 570 | Arg | T;r | Glu | Glu | Asp 575 | Arg |
| | Glu | Ser | Gly | Ala 580 | Trp | Arg | Ile | Arg | Ala 585 | Cys | Ph∈ | Ala | Ala | A≅p 590 | Asn | Glγ |
| 50 | _ | _ | 595 | | | | | 600 | Arg | | | _ | 605 | | | |
| | Tyr | Ala 610 | Pro | Ala | Ile | Met | Pro 615 | Val | Arg | Ala | Phe | G1n 620 | Asp | Val | Tyr | Pro |
| | 625 | | | | | 630 | | | Ser | | 635 | | | | | 640 |
| 55 | | | | | 645 | | | | Ilŝ | 650 | | | | | 655 | |
| | | _ | _ | 660 | | _ | | | Asp 665 | | | | | 670 | | |
| 60 | - | | 675 | _ | | | - | 680 | | | | | 685 | | | |
| | | 690 | - | | | | 695 | - | Ser | | | 700 | | | | |
| e E | 705 | | | | | 710 | | | Gly | | 715 | | | | | 720 |
| 65 | | | | | 725 | | | | Asn | 730 | | | | | 735 | |
| | | | | 740 | | | | | Ly:s 745 | - | | | | 750 | | |
| 70 | - | | 755 | | - | | | 760 | | - | | | 765 | | | |
| | | 770 | | | | | 775 | | Phe - | | | 780 | | | | |
| 7- | 785 | _ | | | | 790 | | | Lyε | | 795 | | | | | 800 |
| <i>7</i> 5 | Gly | Gln | Asp | Pro | Glu | Gly | Gln | Туr | Gly | Thr | Ile | 14et | Leu | Γ.'.ε | Ser | Leu |

WO 99/29870 PCT/AU98/01023

```
805
                                                810
       Glu Arg Val Ala Leu Lys Gly Lys Val Thr Asp Glu Lys Gly Thr Phe
820 825 830
       Asp Glu Thr Phe Ser Gly Lys Val Phe Leu Thr Val Phe Asp Gly Arg
                                      840
       Lys Lys Het Thr Ala Leu Glu Glu Glu Gly Asn Asp Leu Ser Leu Val
850 855 860
       Tyr Tyr Asp Tyr Pro Asn Val Met Tyr Ala Gly Ile Ala Glu Val Lys
865 870 875 880
10
       Asp Gl; Leu Phe Glu Thr Ser Phe Ile Val Pro Lys Asp Val Asn Tyr
885 890 895
       Ser Glu His Glu Gly Arg Ile Asn Leu Tyr Ala Tyr Asn Glu Ser Thr
900 905 910
       Lys Ala Glu Ala Het Gly Val Asp Phe Ser Ile Arg Val Gln Pro Gly 915 920 925
15
       Ile Pro Asp Glu Val Thr Glu Asp Asn Thr Pro Pro Glu Ile Ile Ser
930 935 940
       Cys Phe Leu Asn Asp Ser Thr Phe Arg Ser Gly Asp Glu Val Asn Pro
945 950 955 960
20
       Thr Pro Leu Phe Met Ala Glu Val Phe Asp Leu Asn Gly Ile Asn Ile
965 970 975
       Thr Gly Ser Gly Val Gly His Asp Ile Thr Leu Cys Ile Asp Gly Arg
980 985 990
       Ala Asp Leu Thr Tyr Asn Leu Asn Ala Tyr Phe Thr Ser Ser Ala Thr
995 1000 1005
25
       Asp Ala Gl; Val Gly Thr Ile Leu Phe Het Ile Pro Ala Leu Ala Glu
1010 1015 1020
       Gly Asp His Thr Ala Arg Leu Thr Val Trp Asp Ile Phe Asn Asn Ala
1025 1030 1035
30
       Val His His Asp Phe Ser Phe Arg Val Val Asp Gly Ile Ala Pro Asp 1045 1050 1055
       Val Ala Asp Val Ile Leu Phe Pro Asn Pro Val Arg Glu Ser Ala Thr
1060 1065 1070
       Phe Arg Ile Phe His Asn Arg Pro Gly Ser Asp Leu Asn Val Ala Val
1075 1080 1085
35
       Glu Ile Tyr Asp Phe Thr Gly Arg Leu Val Asn Ser Leu Pro Val Lys
1090 1095 1100
       Thr Tyr Ser Ser Ser Tyr Gly Glu Pro Ile Glu Ile Lys Trp Asp Leu
1105 1110 1115
       Thr Ser Lys Tyr Gly Val Lys Ile Gly Asn Gly Phe Tyr Leu Tyr Arg
1125 1130 1135
40
       Cys Val Val Asn Ser Pro Gly Gly Gln Thr Ala Ser Met Ala Lys Lys
       1140
Het Ile Val Val Gly Gln
45
               1155
       (2) INFORMATION FOR SEQ ID NO:524
             (i) SEQUENCE CHARACTERISTICS:
50
                  (A) LENGTH: 961 amino acids
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) NOLECULE TYPE: protein
55
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Perphyromonas gingivalis
60
           (ix) FEATURE:
                  (A) NAME/KEY: misc feature
                  (B) LOCATION 1...961
65
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:524
      Het Lys Lys Leu Phe Pro Leu Leu Leu Leu Ile Leu Ser Ile Leu Val
                                             10
      Gly Cys Gly Lys Lys Glu Lys His Ser Val Thr Glu Ile Ala Arg Glu
20 25 30
70
       Lys Lys Arg Ile Thr Ala Leu Leu Tyr Glu Lys Glu Leu Pro Thr Asp
              35
                                     40
                                                          45
       Ser Val Lys Gln Leu Tyr Glu Asn Ser Val Gln Asn Lys Asn Leu Val
50 60
75
       Gly Gln Net Leu Phe Ala Ile Glu Val Gly Lys Arg Net Arg Asn Met
```

| | 65 | | | | | 70 | | | | | 75 | | | | | 80 |
|----|------------|-------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Ser | Gln | 47.r | Thr | Asp 85 | Ala | Het | Leu | Туг | His 90 | Gln | Γλε | G17. | Leu | Asn 95 | Ala |
| 5 | Ala | Leu | Asn | Leu 100 | | Asp | Thr | Ile | Val 105 | Ala | Ala | Gln | Ala | Trp 110 | Asn | His |
| | | | 115 | Asp | | | | 120 | Gly | Ala | | | 125 | | | |
| | | 130 | | | | | 135 | | | Glu | | 140 | | | | |
| 10 | Asn 145 | Arg | Pro | Alā | Ile | L;;s 150 | Ala | Arg | Ser | Ala | Ala 155 | Leu | Asn | Glу | Ile | Gl; 160 |
| | | Ile | Asn | Leu | Glu 165 | Leu | Gl ; | Tyr | His | Asp 170 | Glu | Ala | Glu | Lys | Asn 175 | Fhe |
| 15 | Leu | Lys | Ala | Leu 180 | Gln | Gly | Glu | Γλ.ε | Glu 185 | Leu | Asp | Ser | Pro | Leu 190 | G17. | Gln |
| | | | 195 | | | | | 200 | _ | Ile | - | | 205 | _ | | |
| | Tyr | Asp 210 | Lys | Ala | Arg | Thr | Tyr 215 | Phe | Leu | Leu | Ser | Leu 220 | Glu | Gln | Asn | Asn |
| 20 | Het 225 | Alā | Glu | Asn | Leu | Met 230 | G17 | Ile | Gl? | Leu | Cys 235 | Ser | Ile | Asn | Leu | G15 240 |
| | Glu | Val | Asp | Glu | Glu 245 | Lys | Gly | Asp | Tyr | Gln 250 | Lys | Ala | Leu | Gln | Glu 255 | Тук |
| 25 | Ala | Thr | Alâ | Tyr 260 | Lys | Leu | Met | Glu | Gln 265 | Leu | Ser | Asp | Arg | Trp 270 | His | Trp |
| | Leu | Asn | Ser 275 | C):s | Ile | Pro | Het | Ala 280 | Arg | Ile | Asn | Leu | Lys 285 | Gln | Gly | Asn |
| | Glu | Arg 290 | Leu | T;·r | Gln | His | Phe 295 | Ile | Ser | Leu | Ala | Glu 300 | Gly | Thr | Ala | L''.E |
| 30 | 305 | | | | | 310 | | | | Glu | 315 | | | | | 320 |
| | Glu | Asn | Leu | Glu | Arg 325 | Lys | Lys | Glu | T7.r | Lys 330 | Gln | Ala | Leu | Glu | Ala 335 | Phe |
| 35 | C7.2 | Leu | Ser | Lys 340 | Thr | Leu | Ser | Asp | Ser 345 | Het | Ser | Ile | Ala | His 350 | Lys | Val |
| | Ser | Ser | Ile 355 | Gln | Glu | Thr | Arg | Phe 360 | Asn | Tyr | Glu | Arg | Asn 365 | rlæ | Ser | Gln |
| | Lys | Gl u 370 | Leu | Glu | Glu | Ile | Gln 375 | Gln | Val | Ser | Lys | Ala 380 | Γλε | Gln | Glu | Lys |
| 40 | 385 | _ | | | | 390 | | | | ₽ħ€ | 395 | | | | | 400 |
| | | | | | 405 | | | | | Tyr 410 | | | | | 415 | |
| 45 | | | | 420 | | | | | 425 | Leu | | | | 430 | | |
| | - | | 435 | | | | _ | 440 | | Ile | | | 445 | | | |
| | | 450 | - | | | | 455 | | - | Leu | | 460 | | | | |
| 50 | 465 | | | _ | | 470 | | | | Ser | 475 | | | | | 480 |
| | | | | | 485 | | | | | Ile 490 | | | | | 495 | |
| 55 | | | | 500 | | | | | 505 | | | | | 510 | | |
| | | | 515 | | | | | 520 | | Gln | | | 525 | | | |
| CO | | 530 | | | | | 535 | | | Asp | | 540 | | | | |
| 60 | 545 | _ | | | | 550 | | | | Asn | 555 | | | | | 560 |
| | | _ | - | - | 565 | | | | | Ala 570 | _ | | _ | | 575 | |
| 65 | | | | 580 | | | | | 585 | Gl; | | | | 590 | | |
| | | | 595 | | | | | 600 | | Туr | | | 605 | | | |
| 70 | | 610 | | _ | | _ | 615 | _ | | Ser | | 620 | | | | |
| 70 | 625 | | | | | 630 | | | | Glu | 635 | | | | | 640 |
| | Ser | Ala | Phe | Thr | 11e | Ser | Ile | Pro | Thr | Gln 650 | Asn | Gln | Ser | Ser | Ser 655 | Ala |
| 75 | Glu | Ile | Leu | Pro 660 | Trp | Leu | Fro | Ser | Ser 665 | Asp | Asp | Il∈ | Val | Het 670 | Pro | Val |

```
His Ile Ala Pro Asp Asp Ser Pro Thr Ser Pro Met Val Ala Ala Leu
675 680 685
                                       680
        Asn His Arg Phe Glu Asp Glu Arg Pro Thr Ile Leu Val Glu Asp
690 695 700
        Asn Lys Asp Ile Asn Leu Leu Val Lys Leu Leu Leu Cys Asp Arg Tyr 705 710 715 720
  5
        Asn Val Leu Ser Ala Ala Asn Gly Lys Glu Gly Ile Ala Leu Ala Thr
725 730 735
        Glu His Ile Pro Asp Ile Ile Ile Thr Asp Ile Het Het Pro Ile Het
740 745 750
10
        Asp Gl; Ile Glu Net Thr Ile Arg Met Lys Gln Ser Pro Leu Leu Cys
755 760 765
        His Ile Pro Ile Val Ala Leu Thr Ala Lys Ser Thr Glu Gln Asp Arg
770 775 780
        Leu Glu Gly Ile Lys Ser Gly Val Val Ser Tyr Leu Cys Lys Pro Phe
785 790 795 800
15
       Ser Pro Glu Glu Leu Leu Met Arg Ile Glu Gln Leu Leu Lys Asp Arg
805 810 815
       Glu Leu Leu Lys Lys Phe Tyr Het Gln Lys Leu Het Leu Asp Arg Lys 820 825 830
20
       Pro Glu Glu Glu Pro Gln Pro Ile Asp Asp Ser Ser Het Gln Phe Leu
835 840 845
       Leu Ala Ala Lys Asp Ala Val Ser Gly Gly Ile Lys Gln Asn Pro Asp
850 855 860
25
       Phe Ser Ala Gln Asp Leu Ala Glu Lys Het Cys Het Ser Pro Ser Gln 865 870 880
       Leu Asn Arg Lys Leu Thr Ser Val Val Gly Cys Ser Thr Ile Gly Tyr
885 890 895
       Ile Gln Gln Ile Lys Ile Lys Leu Ala Cys Lys Leu Leu Ala Asp Glu
900 905 910
30
       Ser Lys Asn Ile Ser Asp Ile Ser Ile Glu Ala Gly Phe Ser Asp Pro
915 920 925
       Ala Tyr Phe Ser Arg Thr Phe Lys Arg Tyr Het Asn Cys Ser Pro Ser 930 935 940
35
       Gln Tyr Arg Gln Lys Leu Leu Ala Met Pro Gly Ser Asp Lys Glu Thr
40
       (2) INFORMATION FOR SEQ ID NO:525
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 556 amino acids
                   (B) TYPE: amino acid
45
                   (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
50
            (vi) ORIGINAL SOURCE:
                   (A) ORGANISM: Porphyromonas gingivalis
            (ix) FEATURE:
55
                   (A) NAME/KEY: misc feature
                   (B) LOCATION 1...5\overline{5}6
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:525
60
       Net Lys Lys Leu Gln Ala Lys Ala Leu Ile Leu Ala Leu Gly Leu 1 \phantom{\bigg|} 5 \phantom{\bigg|} 10 \phantom{\bigg|} 15
       Phe Gln Leu Pro Ala Ile Ala Gln Thr Gln Met Gln Ala Asp Arg Thr
20 25 30
       Asn Gly Gln Phe Ala Thr Glu Glu Het Gln Arg Ala Phe Gln Glu Thr
65
                                                              45
       Asn Pro Pro Ala Gly Pro Val Arg Ala Ile Ala Glu Tyr Glu Arg Ser 50 60
       Ala Ala Val Leu Val Arg Tyr Pro Phe Gly Ile Pro Met Glu Leu Ile
65 70 80
70
       Lys Glu Leu Ala Lys Asn Asp Lys Val Ile Thr Ile Val Ala Ser Glu 85
       Ser Gln Lys Asn Thr Val Ile Thr Gln Tyr Thr Gln Ser Gly Val Asn
100 105 110
       Leu Ser Asn Cys Asp Phe Ile Ile Ala Lys Thr Asp Ser Tyr Trp Thr
75
```

```
Arg Asp Tyr Thr Gly Trp Phe Ala Net Tyr Asp Thr Asn Lys Val Gly 130 135 140
          Leu Val Asp Phe Ile Tyr Asn Arg Pro Arg Pro Asn Asp Asp Glu Phe 145 150 160
          Pro Lys Tyr Glu Ala Gln Tyr Leu Gly Ile Glu Met Phe Gly Met Lys
165 170 175
          10
          Gln Val Ash Gln Lys Het Lys Asp Tyr Leu Gly Ile Thr His His Asp
210 215 220
          Val Val Gln Asp Pro Asn Gly Glu Tyr Ile Asn His Val Asp Cys Trp
225 230 235 240
          Gly Lys Tyr Leu Ala Pro Asn Lys Ile Leu Ile Arg Lys Val Pro Asp 245 250 255
15
          Asn His Pro Gln His Gln Ala Leu Glu Asp Met Ala Ala Tyr Phe Ala
260 265 270
          Ala Gln Thr Cys Ala Trp Gly Thr Lys Tyr Glu Val Tyr Arg Ala Leu
275 280 285
20
          Ala Thr Asn Glu Gln Pro Tyr Thr Asn Ser Leu Ile Leu Asn Asn Arg
290 295 300
           Val Phe Val Pro Val Asn Gly Pro Ala Ser Val Asp Asn Asp Ala Leu
305 310 315
           Asn Val Tyr Lys Thr Ala Met Pro Gly Tyr Glu Ile Ile Gly Val Lys
325 330 335
25
           Gly Ala Ser Gly Thr Pro Trp Leu Gly Thr Asp Ala Leu His Cys Arg
340 350
           Thr His Glu Val Ala Asp Lys Gly Tyr Leu Tyr Ile Lys His Tyr Pro
355 360 365
30
           Ile Leu Gly Glu Gln Ala Gly Pro Asp Tyr Lys Ile Glu Ala Asp Val
370 375 380
           Val Ser Cys Ala Asn Ala Thr Ile Ser Pro Val Gln Cys Tyr Tyr Arg
385 390 395 400
           lle Asn Gly Ser Gly Ser Phe Lys Ala Ala Asp Het Thr Met Glu Ser
405 410
35
           Thr Gly His Tyr Thr Tyr Ser Phe Thr Gly Leu Asn Lys Asn Asp Lys
420 425 430
           Val Glu T_3r T_3r T_4r T_5r T_7r 
40
           Tyr Pro Phe Ile Gly Glu Pro Asp Pro Phe Lys Phe Thr Cys Met Asn 450 455 460
                                               455
           Glu Thr Asn Thr Cys Thr Val Thr Gly Ala Ala Lys Ala Leu Arg Ala 465 470 480
           Trp Phe Asn Ala Gly Arg Ser Glu Leu Ala Val Ser Val Ser Leu Asn
485 490 495
45
           Ile Ala Gly Thr Tyr Arg Ile Lys Leu Tyr Asn Thr Ala Gly Glu Glu
500 505
           Val Ala Ala Het Thr Lys Glu Leu Val Ala Gly Thr Ser Val Phe Ser
515 520 525
           Het Asp Val Tyr Ser Gln Ala Pro Gly Thr Tyr Val Leu Val Val Glu
530 535 540
            Gly Asn Gly Ile Arg Glu Thr Met Lys Ile Leu Lys
55
            (2) INFORMATION FOR SEQ ID NO:526
                     (i) SEQUENCE CHARACTERISTICS:
                              (A) LENGTH: 428 amino acids
60
                              (B) TYPE: amino acid
                              (D) TOPOLOGY: linear
                   (ii) MOLECULE TYPE: protein
65
                 (iii) HYPOTHETICAL: YES
                   (vi) ORIGINAL SOURCE:
                              (A) ORGANISM: Porphyromonas gingivalis
 70
                   (ix) FEATURE:
                              (A) HAHE/KEY: misc_feature
                              (B) LOCATION 1...428
                    (Mi) SEQUENCE DESCRIPTION: SEQ ID NO:526
 75
```

```
Het Lys Leu Ser Ser Lys Lys Ile Leu Ala Ile Ile Ala Leu Leu Thr
1 5 10
                                              10
       Het Gly His Ala Val Gln Ala Gln Phe Val Pro Ala Pro Thr Thr Gly
                     20
                                           2.5
       Ile Arg Met Ser Val Thr Thr Lys Ala Val Gly Glu Lys Ile Glu 35 40
       Leu Leu Val His Ser Ile Glu Lys Lys Gly Ile Trp Île Asp Leu Asn 50 55
       Gly Asp Ala Thr Tyr Gln Gln Gly Glu Glu Ile Thr Val Phe Asp Glu 65 70 75 80
10
       Ala Tyr His Glu Tyr Thr Ile Gly Thr Gln Thr Leu Thr Ile Tyr Gly
85 90
       Asn Thr Thr Arg Leu Gly Cys Arg Ser Thr Gly Ala Thr Ala Val Asp 100 \hspace{1cm} 105 \hspace{1cm} 110
15
       Val Thr Lys Asn Pro Asn Leu Thr Tyr Leu Ala Cys Pro Lys Asn Asn
115 120 125
       Leu Lys Ser Leu Asp Leu Thr Gln Asn Pro Lys Leu Leu Arg Val Trp
130 135 140
       Cys Asp Ser Asn Glu Ile Glu Ser Leu Asp Leu Ser Gly Asn Pro Ala
145 150 155 160
20
       Leu Ile Ile Leu Gly Cys Asp Arg Asn Lys Leu Thr Glu Leu Lys Thr
165 170 175
       Asp Asn Asn Pro Lys Leu Ala Ser Leu Trp Cys Ser Asp Asn Asn Leu
180 185
25
       Thr Glu Leu Glu Leu Ser Ala Asn Pro Arg Leu Asn Asp Leu Trp Cys
195 200 205
       Phe Gly Asn Arg Ile Thr Lys Leu Asp Leu Ser Ala Asn Pro Leu Leu 210 215 220
       Val Thr Leu Trp Cys Ser Asp Ash Glu Leu Ser Thr Leu Asp Leu Ser
225 230 235
30
       Lys Asn Ser Asp Val Ala Tyr Leu Trp Cys Ser Ser Asn Lys Leu Thr
245 250 255
       Ser Leu Asn Leu Ser Gly Val Lys Gly Leu Ser Val Leu Val Cys His
260 265 270
35
       Ser Asn Gln Ile Ala Gly Glu Glu Met Thr Lys Val Val Asn Ala Leu
275 280 285
       Pro Thr Leu Ser Pro Gly Ala Gly Ala Gln Ser Lys Phe Val Val 290 295 300
       Asp Leu Lys Asp Thr Asp Glu Lys Asn Ile Cys Thr Val Lys Asp Val
305 310 320
40
       Glu Lys Ala Lys Ser Lys Asn Trp Arg Val Phe Asp Phe Asn Gly Asp
325 330 335
       Ser Asp Asn Het Leu Pro Tyr Glu Gly Ser Pro Thr Ser Asn Leu Ala
340 345 350
       Val Asp Ala Pro Thr Val Arg Ile Tyr Pro Asn Pro Val Gly Arg Tyr
355 360 365
45
       Ala Leu Val Glu Ile Pro Glu Ser Leu Leu Gly Gln Glu Ala Ala Leu
370 375 380
       Tyr Asp Net Asn Gly Val Lys Val Tyr Ser Fhe Ala Val Glu Ser Leu
385 390 395
50
       Arg Gln Asn Ile Asp Leu Thr His Leu Pro Asp Gly Thr Tyr Phe Phe
405 410 415
       Arg Leu Asp Asn Tyr Thr Thr Lys Leu Ile Lys Gln
55
       (2) INFORMATION FOR SEQ ID NO:527
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 310 amino acids
60
                   (B) TYPE: amino acid(D) TOPOLOGY: linear
            (ii) NOLECULE TYPE: protein
65
           (iii) HYPOTHETICAL: YES
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
70
            (ix) FEATURE:
                   (A) NAME/KEY: misc_feature
                   (B) LOCATION 1...310
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:527
75
```

```
Het Arg Lys Thr Ile Ile Phe Cys Leu Leu Leu Ala Leu Phe Gly Cys
      Ser Trp Ala Gln Glu Arg Val Asp Glu Lys Val Phe Ser Ala Gly Thr
                                         25
      Ser Ile Phe Arg Gly Ile Leu Glu Lys Val Lys Ala Pro Leu Met Tyr
35 40 45
      Gly Asp Arg Glu Val Trp Gly Het Ala Arg Ala Ser Glu Asp Phe Phe
50 60
      Phe Ile Leu Pro Val Thr Asp Asp Leu Thr Pro Val Leu Phe Tyr Asn 65 70 75 80
10
      Arg Leu Thr Asn Glu Pro Cys Phe Val Ser Asp Gln Gly Ile Thr Glu 85 90 95
      Tyr Phe Lys Phe Ala Gln Glu Gly Asp Tyr Ile Glu Val Glu Gly Ser
100 105 110
      Ser Val Phe Het Ala Asn Leu Leu Tyr Tyr Arg Phe Phe Pro Thr Arg 115 120 125
15
      Ile Thr Ser Tyr Asn Ala Pro Ile Glu Gly Val Val Ser Lys Thr Gly
130 135 140
      Asn Pro Ala Phe Thr Ile Pro Net Leu Pro Gly Val Ser Asp Cys Ile
145 150 155 160
20
      Giu Ile Ser Asn Asn Arg Lys Val Phe Leu Thr Asn Gln Leu Gly Val
165 170
      Val Asn Ile Thr Asp Gly Met Glu Pro Pro Ile Ile Ala Gly Val Ser
180 195 190
25
      Ala Ser Tyr Gly Ser Ser Val Arg Val Tyr Gly His Val Ser Gln Arg
195 200 205
      Trp Asp IIe Ile Gly His Cys Tyr Leu Asp Ile Tyr Prc Thr Asn Cys 210 215 220
      Tyr Pro Leu Ser Thr Lys Pro Val Ala Gly Asp Asp Glu Val Phe Val
225 230 235 249
30
      Lys Gln Gln Gly Arg Gln Ile Glu Ile Asp Ser Asn Ser Pro Ile Val
245 250 255
      Gln Val Val Val Tyr Asp Leu Glu Gly Lys Ser Val Phe Arg Lys Arg
260 265 270
35
      Met Thr Glu Asn Ala Tyr Thr Leu Ser Phe Arg Ala Pro Met Leu Gly 275 280 285
       Phe Net Thr Ile Met Ile Glu Thr Gln Asn Ser Ile Ile Asn Lys Lys
       Leu Asn Val Thr Gln Leu
40
       (2) INFORMATION FOR SEQ ID NO:528
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 405 amino acids
45
                  (B) TYPE: amino acid
                  (D) TOPOLOGY: linear
           (ii) MOLECULE TYPE: protein
50
          (iii) HYPOTHETICAL: YES
           (vi) ORIGINAL SOURCE:
                 (A) ORGANISM: Porphyromonas gingivalis
55
           (ix) FEATURE:
                 (A) NAME/KEY: misc feature
                  (B) LOCATION 1...405
60
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:528
      Het Lys Lys Thr Thr Ile Ile Ser Leu lle Val Phe Gly Ala Phe Phe
                                              10
      Ala Ala Val Gly Gln Thr Lys Asp Asn Ser Ser Tyr Lys Pro Phe Ser
65
                                         25
      Lys Glu Asp Ile Ala Gly Gly Val Tyr Ser Leu Pro Thr Gln Asn Arg
35 40
      Ala Gln Lys Asp Asn Ala Glu Trp Leu Leu Thr Ala Thr Val Ser Thr 50 60
70
      Asn Gln Ser Ala Asp Thr His Phe Ile Phe Asp Glu Asn Asn Arg Tyr
65 70 75 80
      Ile Ala Arg Asp Ile Lys Ala Asn Gl; Val Arg Lys Ser Thr Asp Ser
                                             90
                       85
      Ile Tyr Tyr Asp Ala Asn Gly Arg Ile Ser His Val Asp Leu Tyr Ile
75
```

WO 99/29870 PCT/AU98/01023

```
Ser Phe Ser Gly Gly Glu Pro Ala Leu Asp Thr Arg Phe Lys Tyr Thr
115 120 125
        115 120 125
Tyr Asp Asp Glu Gly Lys Met Thr Val Arg Glu Val Phe Met Leu Val
            130
                                   135
        Met Asp Pro Asn Thr Pro Ile Ser Arg Leu Glu Tyr His Tyr Asp Ala
                             150
                                                    155
        Gln Gly Arg Leu Thr His Trp Ile Ser Phe Ala Phe Gly Ala Glu Ser
165 170 175
        Gln Lys Asn Thr Tyr His Tyr Asn Glu Lys Gly Leu Leu Val Ser Glu
180 185 190
10
        Val Leu Ser Asn Ala Met Gly Thr Thr Tyr Ser Asp Thr Gly Lys Thr
195 200 205
        Glu Tyr Ser Tyr Asp Asp Ala Asp Asn Het Val Lys Ala Glu Tyr Phe
210 215 220
15
        Val Val Gln Gln Gly Lys Ala Trp Gln Val Leu Lys Arg Glu Glu Tyr
225 230 235 240
        Thr Tyr Glu Asp Asn Ile Cys Ile Gln Tyr Leu Ala Ile Asn Gly Thr
245 250 255
        Asp Thr Lys Val Tyr Lys Arg Asp Ile Glu Ser Asp Lys Ser Ile Ser 260 265 270
20
                                              265
                                                                      270
        Ala Asn Val Ile Asp Ile Pro Ser Met Pro Glu Gln Thr Trp Pro Asn
275 280 285

    Het
    Tyr
    Gly
    Phe
    Asn
    Ala
    Lys
    Arg
    Leu
    Lys
    Glu
    Thr
    Tyr
    Ser
    Ser
    Tyr

    290
    295
    300

    Glu
    Gly
    Asp
    Val
    Ala
    Thr
    Fro
    Ile
    Phe
    Asp
    Tyr
    Ile
    Tyr
    Thr
    Tyr
    Lys

    305
    310
    315
    320

25
        Ala Leu Thr Ser Het Ala Thr Pro Ser Thr Glu Ala Gln Val Ala Val
325 330 335
       Tyr Leu Asn Pro Ser Thr Asp Arg Leu Val Ile Leu Ala Asn Gly Ile 340 \hspace{1cm} 345 \hspace{1cm} 350
30
        Thr His Leu Ser Met Tyr Asp Leu Gln Gly Lys Leu Ile Arg Asp Cys 355 360
                                                            365
        Ala Leu Ser Gly Asp Lys Val Glu Met Gly Val Gly Ser Leu Thr Lys
370 375 380
                                                           380
35
        Gly Thr Tyr Leu Leu Lys Val Asn Thr Asp Gln Gly Ala Phe Val Arg
                                                         395
       Lys Val Val Ile Arg
                           405
40
        (2) INFORMATION FOR SEQ ID NO:529
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 2037 base pairs
                    (B) TYPE: nucleic acid
45
                    (C) STRANDEDNESS: double
                    (D) TOPOLOGY: circular
             (ii) HOLECULE TYPE: DNA (genomic)
50
            (iii) HYPOTHETICAL: NO
            (iv) ANTI-SENSE: NO
            (vi) ORIGINAL SOURCE:
55
                    (A) ORGANISH: Porphyromonas gingivalis
             (ix) FEATURE:
                    (A) MAME/KEY: misc feature
                    (B) LOCATION 1...2037
60
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:529
       GTAATGGACT ACAAACTCAC TTCTCGATTC AAGCCCACGG GCGACCAGCC GGAAGCCATT
       CGCCAACTCG TACAGGGCAT CAACGAAGGG ATGCCGGCTC AGACGCTGCT CGGCGTAACG
GGTTCGGGCA AAACTTTTAC GGTGGCTAAC GTGGTGCGG CGGTCAATCG TCCGACCCTT
                                                                                               120
65
                                                                                               180
       GTCCTGAGTC ACAACAAGAC CTTGGCAGCA CAGCTATACG GAGAGTTCAA AGCCTTCTTC
CCCGAGAATG CGGTGGAGTA TTTCGTCAGC TACTACGACT ACTATCAGCC CGAGGCCTAC
                                                                                               240
       CTCCCCGTCA CAGACACCTA TATCGAAAAG GACATGGCCA TCAACGCGGA GATCGAAAAA
                                                                                               360
       CTGCGATTGA GGGCCACGGC TTCGCTCCTG TCAGGGCGGA AAGATGTGCT TGTGGTCAGC
                                                                                               420
70
       TCCGTATCCT GTCTCTACGG TATGGCCAAT CCTGAAGCTT TTTCCGAAAA GGTGATCAGC
                                                                                               480
       CTGCACACGG GACAAAGGGC AGACAGGGAT CATTTTATCC GCCTGCTGGT AGAGAGCTAC
                                                                                               540
       TACACGAACA ATAAAGTAGA GTTCGAGAGC GGCAACTTCC GTGTCAAAGG CGACAGCGTG
                                                                                               600
       GACATATTCC CCGCCGTAGA AGGTTATGAC GGCGTGGCAT ACAGGGTGGA GTTTTGGGAT
                                                                                               660
       GGAGAGGTCG AGCGGCTGAG TACCTTCGAT CCGCGAACGG GACGGGAATA CGGCCTGCTG
75
       TOGGAGOTGA AGATATATOO GGOCAATOTO TTOGTGACGA CTAAGGAGCA GGTGGATOGG
```

WO 99/29870 PCT/AU98/01023

```
GCAGTAGGGA AAATCGATGT GGATCTGGGC GCACAGGTCG ATTTTCTGAA AGAAATCGGC
                                                                                       840
       AAACCATATG AAGCCAAACG CTTGTATGAG CGGGTCACGT ATGACTTGGA AATGATCCGT
                                                                                       900
                                                                                       960
       GAGTTGGGTT ATTGTTCCGG TATAGAGAAC TATTCGCGCT ACTTCGACGG CCGTGACGCG
       GGCGAACGTC CTTTCTGTCT GTTGGATTAT TTCCCGGAGG ATTTCCTGTT GGTCATAGAC
                                                                                      1020
 5
       GANAGCCATG TRACGATACC GCAGATACGT GCCATGTACG GAGGCGATCG TTCGCGCAAG
                                                                                      1080
       GAGAATCTGG TCGAATACGG ATTCCGCCTG CCTGCCGCTC TCGACAATCG GCCGCTTCGC
                                                                                      1140
       TTCGACGAGT TCGAAGCTCT CACCCCCGG ACCCTTTATA TCAGTGCCAC GCCTGCCGAC
                                                                                      1200
       TATGAGCTGA ACAGAAGCGA AGGCGTGATC GTCGAGCAGC TGATCCGTCC GACCGGACTG
                                                                                      1260
       CTGGATCCCA TCATCGACGT CAAGCCGACG GCAAACCAAG TGGACGATCT GATGGAGGAG
                                                                                      1320
10
       ATAGCACGCT GCATCGAAAA GAAAGAGCGC GTACTGGTAA CGACCCTGAC CAAACGTATG
                                                                                      1389
       GCAGAGGAGC TTAGCGAATA CCTGCTACGC CACGGTATCA GCACCGGCTA CATACACAGC
      GATGTGGACA CGCTGGAGCG TGTGCGTATC ATGGAAGACC TGCGCAAGGG GGTCTACGAT
GCACTCATCG GGGTGAATCT GCTCCGCGAA GGATTGGACT TGCCGGAAGT TTCGCTTGTG
                                                                                      1500
                                                                                      1560
      GCTATTCTGG ATGCGGATAA GGAAGGATTC CTGCGCTCGC ATCGTTCGCT CACGCAGACT
GCAGGACGTG CCGCCCGGCA CATTCATGGG CGTGTCATCT TCTACGCGGA CAAGATCACC
                                                                                      1620
15
                                                                                      1680
      GACAGTATGC AGCTCACCAT GGACGAGACT GCACGCCGAC GCGCAAAGCA ACTGGCCTAC AACGAAGCG ACGGCATCAC CCCCCAACAG ATAGTGAAGA ACAGTGCTGC CATTTGGGGA
                                                                                      1740
                                                                                      1800
       GAAGGCGATG TGTCGGCCTT GCAATCCGAT ACAGAATCCG GTGCGTACAT AGAAGAGAGC
                                                                                      1860
       AGCATGGTGG CTGCCGATCC TTTGGCCGAC TATCTGAGCA AACCCAAGCT GGAAGCACTC
                                                                                      1920
20
       ATTGCTTCGA CCAAGAAGCA AATGCTGGCA GCAGCCAAAG AGCTGGACTT TCTGGAAGCG
                                                                                      1980
       GCACGACTTC GGGACGAAGC CGCACGATTG GAAAAGAAAC TGGAGCAACT CACAGCC
                                                                                      2037
       (2) INFORMATION FOR SEQ ID NO:530
25
            (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 2034 base pairs
                  (B) TYPE: nucleic acid
                   (C) STRAHDEDHESS: double
30
                  (D) TOPOLOGY: circular
           (ii) MOLECULE TYPE: DMA (genomic)
          (iii) HYPOTHETICAL: NO
35
           (iv) ANTI-SENSE: NO
           (vi) ORIGINAL SOURCE:
                  (A) ORGANISM: Porphyromonas gingivalis
40
           (ix) FEATURE:
                  (A) NAME/KEY: misc_feature
                  (B) LOCATION 1...2034
45
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:530
       ATGGACTACA AACTCACTTC TCGATTCAAG CCCACGGCGG ACCAGCCGGA AGCCATTCGC
       CAACTOSTAC AGGGCATCAA CGAAGGGATG CCGGCTCAGA CGCTGCTCGG CGTAACGGGT
                                                                                       120
       TCGGGCAAAA CTTTTACGGT GGCTAACGTG GTGGCGGCGG TCAATCGTCC GACCCTTGTC
                                                                                       180
50
       CTGAGTCACA ACAAGACCTT GGCAGCACAG CTATACGGAG AGTTCAAAGC CTTCTTCCCC
                                                                                       240
       GAGAATGCGG TGGAGTATTT CGTCAGCTAC TACGACTACT ATCAGCCCGA GGCCTACCTC
CCCGTCACAG ACACCTATAT CGAAAAGGAC ATGGCCATCA ACGCGGAGAT CGAAAAACTG
                                                                                       300
                                                                                       360
      CGATTGAGGG CCACGGCTTC GCTCCTGTCA GGGCGGAAAG ATGTGCTTGT GGTCAGCTCC GTATCCTGTC TCTACGGTAT GGCCAATCCT GAAGCTTTTT CCGAAAAGGT GATCAGCCTG
                                                                                       420
                                                                                       480
       CACACGGGAC AAAGGGCAGA CAGGGATCAT TTTATCCGCC TGCTGGTAGA GAGCTACTAC ACGAACAATA AAGTAGAGTT CGAGAGCGGC AACTTCCGTG TCAAAGGCGA CAGCGTGGAC
55
                                                                                       540
                                                                                       600
                                                                                       660
       ATATTCCCCG CCGTAGAAGG TTATGACGGC GTGGCATACA GGGTGGAGTT TTGGGATGGA
       GAGGTCGAGC GGCTGAGTAC CTTCGATCCG CGAACGGGAC GGGAATACGG CCTGCTGTCG
                                                                                       720
       GAGCTGAAGA TATATCCGGC CAATCTCTTC GTGACGACTA AGGAGCAGGT GGATCGGGCA
                                                                                       780
60
       GTAGGGAAAA TCGATGTGGA TCTGGGCGCA CAGGTCGATT TTCTGAAAGA AATCGGCAAA
                                                                                       840
       CCATATGAGG CCAAACGCTT GTATGAGCGG GTCACGTATG ACTTGGAAAT GATCCGTGAG
                                                                                       900
       TTGGGTTATT GTTCCGGTAT AGAGAACTAT TCGCGCTACT TCGACGGCCG TGACGCGGGC
GAACGTCCTT TCTGTCTGTT GGATTATTTC CCGGAGGATT TCCTGTTGGT CATAGACGAA
                                                                                      1020
       AGCCATGTAA CGATACCGCA GATACGTGCC ATGTACGGAG GCGATCGTTC GCGCAAGGAG
                                                                                      1080
65
       1140
       GACGAGTTCG AAGCTCTCAC CCCCGGACC CTTTATATCA GTGCCACGCC TGCCGACTAT
                                                                                      1200
       GAGCTGAACA GAAGCGAAGG CETGATCGTC GAGCAGCTGA TCCGTCCGAC CGGACTGCTG
                                                                                      1260
       GATCCCATCA TCGACGTCAA GCCGACGGCA AACCAAGTGG ACGATCTGAT GGAGGAGATA
                                                                                      1320
       GCACGCTGCA TCGAAAAGAA AGAGCGCGTA CTGGTAACGA CCCTGACCAA ACGTATGGCA
                                                                                      1380
70
       GAGGAGCTTA GCGAATACCT GCTACGCCAC GGTATCAGCA CCGGCTACAT ACACAGCGAT
GTGGACACGC TGGAGCGTGT GCGTATCATG GAAGACCTGC GCAAGGGGGT CTACGATGCA
                                                                                      1440
                                                                                      1500
       CTCATCGGGG TGAATCTGCT CCGCGAAGGA TTGGACTTGC CGGAAGTTTC GCTTGTGGCT
                                                                                      1560
       ATTCTGGATG CGGATAAGGA AGGATTCCTG CGCTCGCATC GTTCGCTCAC GCAGACTGCA
                                                                                      1620
       GGACGTGCCG CCCGGCACAT TCATGGGCGT GTCATCTTCT ACGCGGACAA GATCACCGAC
                                                                                      1690
75
       AGTATGCAGC TCACCATGGA CGAGACTGCA CGCCGACGCG CAAAGCAACT GGCCTACAAC
                                                                                      1740
```

PCT/AU98/01023

| 5 | GGCGATG ATGGTGG GCTTCGA | ACG GCATO TGT CGGCO CTG CCGAO CCA AGAAO GGG ACGA | CTTGCA A TCCTTT G GCAAAT G | TCCG/ GCCG/ CTGG(| ATACI ACTAI CAGCI | A GAZ CTC A GCC | ATCC: SAGCA CAAAC | GTG NAAC SAGC | CGTA CCAA TGGA | ACATA AGCTO ACTTI | AGA A GGA A CT (| AGAGA AGCAC EGAAC | AGCAGC CTCATT |
|------------|-------------------------------|--|--|-------------------------|-------------------------|-----------------------|-------------------------|---------------------|----------------------|-------------------------|------------------------|-------------------------|------------------|
| | (2) IHF | ORHATION | FOR SEQ | ID! | 10:53 | 31 | | | | | | | |
| 10 | (i | (B) T | CE CHARA ENGTH: 6 YPE: ami OPOLOGY: | 79 ar no ac | mino cid | | de | | | | | | |
| 15 | (ii |) HOLECU | LE TYPE: | prot | ein | | | | | | | | |
| | (iii |) HYPOTH | ETICAL: | YES | | | | | | | | | |
| 20 | (vi |) ORIGINA (A) O | AL SOURC RGANISH: | | ohyro | omona | es đị | ngiv | alis | Ē | | | |
| 25 | (ix | | E: AHE/KEY: OCATIOH | | | ture | е | | | | | | |
| | (xi |) SEQUEII | CE DESCR | IPTIC | on: s | SEQ] | (D 110 | 0:531 | - | | | | |
| | Val Het 1 | Asp Tyr | Lys Leu 5 | Thr | Ser | Arg | Ph∈ 10 | Lys | Pro | Thr | Glγ | Asp 15 | Gln |
| 30 | | Ala Ile 20 | Arg Gln | Leu | Val | G1n 25 | | Ile | Asn | Glu | Gly 30 | | Pro |
| | Ala Gln | Thr Leu | Leu Gly | Val | Thr 40 | | Ser | Gly | Lys | Thr 45 | | Thr | Val |
| 35 | Ala Asn 50 | Val Val | Ala Ala | Val 55 | | Arg | Pro | Thr | Leu 60 | | Leu | Ser | His |
| | | Thr Leu | Ala Ala 70 | | Leu | Tyr | Gly | Glu 75 | | ГЛЗ | Ala | Phe | Phe 80 |
| | Pro Glu | Asn Ala | | Tyr | Phe | Val | Ser 90 | | Tyr | Asp | Tyr | Tyr 95 | |
| 4 0 | Pro Glu | Ala Tyr 100 | Leu Pro | Val | Thr | Asp 105 | | Tyr | Ile | Glu | Lys 110 | Asp | lfet |
| | Ala Ile | Asn Ala 115 | | Glu | Lys 120 | | Arg | Leu | Arg | Ala 125 | | Ala | Ser |
| 45 | Leu Leu 130 | Ser Gly | Arg Lys | Asp 135 | | Leu | Val | Val | Ser 140 | | Val | Ser | Cys |
| | Leu Tyr 145 | Gly Met | Ala Asn 150 | Pro | Glu | Ala | Pine | Ser 155 | | Lys | Val | Ile | Ser 160 |
| | Leu His | Thr Gly | Gln Arg 165 | Ala | Asp | Arg | Asp 170 | | Phe | Ile | Arg | Leu 175 | |
| 50 | Val Glu | Ser Tyr 180 | Tyr Thr | Asn | Asn | Lys 185 | | Glu | Phe | Glu | Ser 190 | | Asn |
| | Phe Arg | Val Lys 195 | | Ser | Val 200 | | lle | Phe | Pro | Ala 205 | | Glu | Gly |
| 55 | Tyr Asp 210 | Gly Val | Ala Tyr | Arg 215 | | Glu | Phe | Trp | Asp 220 | | Glu | Val | Glu |
| | | Ser Thr | Phe Asp 230 | Pro | Arg | Thr | Glу | Arg 235 | | Tyr | Gly | Leu | Leu 240 |
| | | Leu Lys | | | Alā | Asn | Leu 250 | | Val | Thr | Thr | Lys 255 | |
| 60 | Gln Val | Asp Arg 260 | Ala Val | Gly | Lys | Ile 265 | | Val | Asp | Leu | Gly 270 | | Gln |
| | Val Asp | Phe Leu 275 | | Ile | Gly 280 | | Pro | Tyr | Glu | Ala 285 | | Arg | Leu |
| 65 | Tyr Glu 290 | Arg Val | Thr Tyr | Asp 295 | | Glu | Het | lle | Arg 300 | | Leu | Gly | Tyr |
| | | Gly Ile | Glu Asn 310 | Tyr | Ser | Arg | Tyr | Phe | | Gl y | Arg | Asp | Ala 320 |
| | | Arg Pro | | | Leu | Asp | Tyr 330 | | Pro | Glu | Asp | Phe 335 | |
| 70 | Leu Val | Ile Asp 340 | Glu Ser | His | Val | Thr 345 | | Pro | Gln | Ile | Arg 350 | | Het |
| | Tyr Gly | Gly Asp 355 | | Arg | Lys 360 | | Asn | Leu | Vāl | Glu 365 | | Gly | Phe |
| 75 | Arg Leu 370 | Pro Ala | Ala Leu | Asp 375 | | Arg | Pro | Leu | Arg 380 | | Asp | Glu | Phe |

385

489/490

```
395
        Tyr Glu Leu Asn Arg Ser Glu Gly Val Ile Val Glu Gln Leu Ile Arg
                       405
                                                410
       Pro Thr Gly Leu Leu Asp Pro Ile Ile Asp Val Lys Pro Thr Ala Asn
420 425 430
        Gln Val Asp Asp Leu Met Glu Glu Ile Ala Arg Cys Ile Glu Lys Lys
435 440 445
                                                            445
        Glu Arg Val Leu Val Thr Thr Leu Thr Lys Arg Met Ala Glu Glu Leu
450 455 460
 10
        Ser Glu Tyr Leu Leu Arg His Gly Ile Ser Thr Gly Tyr Ile His Ser
465 470 475
       Asp Val Asp Thr Leu Glu Arg Val Arg Ile Net Glu Asp Leu Arg Lys
485
490
495
Gly Val Tyr Asp Ala Leu Ile Gly Val Asn Leu Leu Arg Glu Gly Leu
500
500
 15
       Asp Leu Pro Glu Val Ser Leu Val Ala Ile Leu Asp Ala Asp Lys Glu
515 520 525
       Gly Phe Leu Arg Ser His Arg Ser Leu Thr Gln Thr Ala Gly Arg Ala
530 535
20
       Ala Arg His Ile His Gly Arg Val Ile Phe Tyr Ala Asp Lys Ile Thr
545 550 560
       Asp Ser Net Gln Leu Thr Het Asp Glu Thr Ala Arg Arg Arg Ala Lys
565 570
25
       Gln Leu Ala Tyr Asn Glu Ala His Gly Ile Thr Pro Gln Gln Ile Val
       Lys Asn Ser Ala Ala Ile Trp Gly Glu Gly Asp Val Ser Ala Leu Gln
595 600 605
       Ser Asp Thr Glu Ser Gly Ala Tyr Ile Glu Glu Ser Ser Met Val Ala 610 620
30
       Ala Asp Pro Leu Ala Asp Tyr Leu Ser Lys Pro Lys Leu Glu Ala Leu 625 630 630
       Ile Ala Ser Thr Lys Lys Gln Met Leu Ala Ala Ala Lys Glu Leu Asp
645 650 655
35
       Phe Leu Glu Ala Ala Arg Leu Arg Asp Glu Ala Ala Arg Leu Glu Lys
660 665 670
                                          665
       Lys Leu Glu Gln Leu Thr Ala
40
       (2) INFORMATION FOR SEQ ID NO:532
             (i) SEQUENCE CHARACTERISTICS:
                  (A) LENGTH: 678 amino acids
                   (B) TYPE: amino acid
45
                  (D) TOPOLOGY: linear
            (ii) MOLECULE TYPE: protein
           (iii) HYPOTHETICAL: YES
50
            (vi) ORIGINAL SOURCE:
                  (A) ORGANISH: Porphyromonas gingivalis
           (ix) FEATURE:
55
                  (A) NAME/KET: misc_feature
                  (B) LOCATION 1...678
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:532
60
       Het Asp Tyr Lys Leu Thr Ser Arg Phe Lys Pro Thr Gly Asp Gln Pro
       Glu Ala 11e Arg Gln Leu Val Gln Gly I1e Asn Glu Gly Het Pro Ala
20 25 30
       Gln Thr Leu Leu Gly Val Thr Gly Ser Gly Lys Thr Phe Thr Val Ala 35 40
65
       Asn Val Val Ala Ala Val Asn Arg Pro Thr Leu Val Leu Ser His Asn
50 55
       Lys Thr Leu Ala Ala Gln Leu Tyr Gly Glu Phe Lys Ala Phe Phe Pro 65 \phantom{+}70\phantom{0}
70
       Glu Asn Ala Val Glu Tyr Fhe Val Ser Tyr Tyr Asp Tyr Tyr Gln Pro
85 90
       Glu Ala Tyr Leu Pro Val Thr Asp Thr Tyr Ile Glu Lys Asp Met Ala
100 105 110
       Ile Asn Ala Glu Ile Glu Lys Leu Arg Leu Arg Ala Thr Ala Ser Leu
75
```

Glu Ala Leu Thr Pro Arg Thr Leu Tyr Ile Ser Ala Thr Pro Aia Asp

390

PCT/AU98/01023

| | Leu | | Gly | Arg | Lys | Asp | | Leu | Val | Val | Ser | | Val | Ser | Суѕ | Leu |
|-----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Tyr | G1 y | Het | Ala | Asn | Pro | 135 Glu | Ala | Phe | Ser | | 140 Lys | Vāl | Ile | Ser | |
| 5 | 145 His | Thr | Gly | Gln | Arg | 150 Ala | Asp | Arg | Asp | His | 155 Phe | Ile | Arg | Leu | | 160 Val |
| | Glu | Ser | Туr | Tyr 180 | 165 Thr | Asn | Asn | Lys | | 170 Glu | Phe | Glu | Ser | | 175 Asn | Phe |
| 10 | Arg | Val | Lys 195 | | Asp | Ser | Val | | 185 Ile | Phe | Pro | Ala | | 190 Glu | Gly | туг |
| 10 | Asp | Gly 210 | | Ala | Туг | Arg | Val 215 | 200 Glu | Phe | Trp | Asp | Gly 220 | 205 Glu | Val | Glu | Arg |
| | Leu 225 | | Thr | Phe | Asp | Pro 230 | | Thr | Gly | Arg | Glu 235 | | Gly | Leu | Leu | Ser 240 |
| 15 | | Leu | Lys | Ile | Tyr 245 | | Ala | Asn | Leu | Phe 250 | | Thr | Thr | Lys | Glu 255 | |
| | Val | Asp | Arg | Ala 260 | | Gly | Lys | Пе | Asp 265 | Val | Asp | Leu | Gly | Ala 270 | | Val |
| 20 | Asp | Phe | Leu 275 | | Glu | Ile | Glу | Lys 280 | | Tyr | Glu | Ala | Lys 285 | | Leu | Туг |
| | Glu | Arg 290 | Val | Thr | Tyr | Asp | Leu 295 | | Net | Il∈ | Arg | Glu 300 | | GJ2. | Tyr | Суѕ |
| | Ser 305 | Gly | Ile | Glu | Asn | Tyr 310 | Ser | Arg | Tyr | Ph∈ | Asp 315 | Gly | Arg | Asp | Ala | Gly 320 |
| 25 | | | | | 325 | | | | | Phe 330 | | | | | 335 | |
| | | | | 340 | | | | | 345 | Pro | | | | 350 | | |
| 30 | | | 355 | | | | | 360 | | Leu | | | 365 | | | |
| | | 370 | | | | | 375 | | | Leu | | 380 | | | | |
| or. | 385 | | | | | 390 | | | | Ser | 395 | | | | | 400 |
| 35 | | | | | 405 | | | | | Val 410 | | | | | 415 | |
| | | | | 420 | | | | | 425 | Vāl | | | | 430 | | |
| 40 | | | 435 | | | | | 440 | | Arg | | | 445 | | | |
| | | 450 | | | | | 455 | | | Arg | | 460 | | | | |
| 45 | 465 | | | | | 470 | | | | Thr | 475 | | | | | 480 |
| 10 | | | | | 485 | | | | | Het 490 | | | | | 495 | |
| | | | | 500 | | | | | 505 | Leu | | | | 510 | | |
| 50 | | | 515 | | | | | 520 | | Gln | | | 525 | | | |
| | | 530 | | | | | 535 | | | Tyr | | 540 | | | | |
| 55 | 545 | | | | | 550 | | | | Ala | 555 | | | | | 560 |
| | | | | | 565 | | | | | 570 Thr | | | | | 575 | |
| | | | | 580 | | | | | 585 | Asp | | | | 590 | | |
| 60 | | | 595 | | | | | 600 | | Glu | | | 605 | | | |
| | | 610 | | | | | 615 | | | Pro | | 620 | | | | |
| 65 | 625 | | | | | 630 | | | | Ala | 635 | | | | | 640 |
| | | | | | 645 | | | | | 650 Ala | | | | | 655 | |
| 70 | Leu | | Gln | 660 | | | - | - | 665 | | | - | | 670 | - | - |
| 70 | | | 675 | | | | | | | | | | | | | |

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU 98/01023

| | | | |
|--|---|--|---|
| Α. | CLASSIFICATION OF SUBJECT MATTER | | |
| Int Cl ⁶ : | C12N 15/31; C07K 14/195; A61K 38/00, 38/16 | , 39/00 | |
| According to | International Patent Classification (IPC) or to both | h national classification and IPC | |
| В. | FIELDS SEARCHED | | |
| | umentation searched (classification system followed by a ic Database box below | classification symbols) | |
| | n searched other than minimum documentation to the extic Database box below | tent that such documents are included in | the fields searched |
| Electronic data Medline CA WPAT | base consulted during the international search (name of porphyromonas gingivalis, bacterial protestions) porphyromonas gingivalis, genes microbia porphoryomonas or porphyromonas gingi 528, 340, 341 | eins, peptides, bacterial vaccines al, antigens | , in the second |
| C. | DOCUMENTS CONSIDERED TO BE RELEVANT | Γ | |
| Category* | Citation of document, with indication, where ap | propriate, of the relevant passages | Relevant to claim No. |
| A | Derwent abstract 95-3513214/45 WO 95/26 5 October 1995. | 6404 (MEITO SANGYO KK) | all |
| A | Derwent abstract 95-147309/19 WO 95/091 KK) 6 April 1995. | 181 (KYOWA HAKKO KOGYO | ali |
| | Further documents are listed in the continuation of Box C | X See patent family an | nex |
| "A" docum not co "E" earlier the in docum or wh anoth "O" docum exhibi | al categories of cited documents: "The nent defining the general state of the art which is possidered to be of particular relevance repolication or patent but published on or after ternational filing date ment which may throw doubts on priority claim(s) aich is cited to establish the publication date of ere citation or other special reason (as specified) ment referring to an oral disclosure, use, ition or other means ment published prior to the international filing but later than the priority date claimed | priority date and not in conflict with understand the principle or theory us document of particular relevance; the be considered novel or cannot be consinuentive step when the document is document of particular relevance; the be considered to involve an inventive combined with one or more other succombination being obvious to a personal conflict with the considered to involve an inventive combination being obvious to a personal conflict with the conflict wit | the application but cited to aderlying the invention e claimed invention cannot asidered to involve an taken alone e claimed invention cannot e step when the document is ch documents, such on skilled in the art |
| | ual completion of the international search | Date of mailing of the international sear 2 8 JAN 1999. | ch report |
| AUSTRALIAN PO BOX 200 WODEN ACT AUSTRALIA | | Authorized officer PHILIPPA WYRDEMAN Talacher No. (22) (222 2554 | |
| I Facsimile No: | (02) 6285 3929 | Telephone No.: (02) 6283 2554 | |

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 98/01023

| Box 1 | Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet) |
|------------|--|
| This inter | rnational search report has not been established in respect of certain claims under Article 17(2)(a) for the following |
| 1. | Claims Nos.: |
| | because they relate to subject matter not required to be searched by this Authority, namely: |
| | |
| | |
| <u>2.</u> | X Claims Nos.: 1 to 34 in part |
| | X Claims Nos.: 1 to 34 in part because they relate to parts of the international application that do not comply with the prescribed requirements to |
| | such an extent that no meaningful international search can be carried out, specifically: |
| and the | |
| | ns are directed to a very large number of polypeptides and nucleotides such that it is economically unfeasible m a full and comprehensive search on all the claimed sequences. |
| - | <u> </u> |
| 3. | Claims Nos.: |
| | because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a) |
| Box II | Observations where unity of invention is lacking (Continuation of item 2 of first sheet) |
| This Inter | national Searching Authority found multiple inventions in this international application, as follows: |
| | and the same of th |
| | |
| | |
| | |
| 1. | As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims |
| 2. | As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee. |
| 3. | As only some of the required additional search fees were timely paid by the applicant, this international search |
| | report covers only those claims for which fees were paid, specifically claims Nos.: |
| | |
| | |
| 4. | No required additional search fees were timely paid by the applicant. Consequently, this international search |
| | report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: |
| | |
| | |
| | |
| Remark o | The additional search fees were accompanied by the applicant's protest. |
| | No protest accompanied the payment of additional search fees. |
| | |

INTERNATIONAL SEARCH REPORT

Information on patent family members

| International application | No |
|---------------------------|----|
| PCT/AU 98/01023 | |

END OF ANNEX

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

| Patent Doo | cument Cited in Search Report | Patent Family Member | | | | | | |
|------------|----------------------------------|----------------------|--------|----|---------|--|--|--|
| WO | 95/26404 | EP | 753577 | | | | | |
| WO | 95/09181 | EP | 726276 | JP | 7097395 | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |

Form PCT/ISA/210 (extra sheet) (July 1998) copbko

